PB-217 667

OPERATIONS RESEARCH IN WATER QUALITY MANAGEMENT

Harold A. Thomas, Jr., et al

Harvard University
Cambridge, Massachusetts

15 February 1963

DISTRIBUTED BY:

NTIS
National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

PH86-62-140

Contr PH 86-62-140

PB 217 667

OPERATIONS RESEARCH IN WATER QUALITY MANAGEMENT

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
U S Department of Commerce
Springfield VA 22151

harvard
water
resources
group

179

HARVARD UNIVERSITY
Division of Engineering and Applied Physics

OPERATIONS RESEARCH IN WATER QUALITY MANAGEMENT

Final Report
to
Bureau of State Services
Division of Water Supply and Pollution Control
Public Health Service
Department of Health, Education and Welfare

Contract PH86-62-140
February 15, 1963
Cambridge, Massachusetts

Harold A. Thomas, Jr.,
Professor Civil and Sanitary
Engineering

Robert P. Burden,
Assistant to the Dean
Division of Engineering and
Applied Physics

I

TABLE OF CONTENTS

Preface . . . . i to ii

List of Figures . . . . iii to v

List of Tables vi

Introduction . . . . 1 to 8

Chapter 1 - Statistical Analysis of Reservoir Storage-Yield Relations . . . . 1-1 to 1-21

Chapter 2 - Computer Simulation of Multipurpose Water Resource Systems . . . . 2-1 to 2-39

Chapter 3 - Queueing Model for Pollution Transport in Streams . . . . 3-1 to 3-18

Appendices to Chapter 3 . . . . 3-19 to 3-23

Appendix A - Standard Notation

Appendix B - General Discussion of Queueing Theory

Appendix C - Finite Markov Chains

Chapter 4 - Application of Decision Theory to Sanitary Engineering Design Problems . . . . 4-1 to 4-16

Chapter 5 - Mathematical Model for Determining the Optimal Sizes of Water Treatment Plants . . . . 5-1 to 5-21

Chapter 6 - A Method of Optimizing Designs of Water Quality Management Systems . . . . 6-1 to 6-14

Chapter 7 - Imputed Benefits of Municipal Water Supply . . . . 7-1 to 7-5

Glossary . . . . G-1 to G-3

PREFACE

Traditionally, the design of water resource systems has fallen within the purview and responsibility of hydraulic and sanitary engineers. Supply of water and disposal of waste water in conjunction with other water uses such as hydropower have, at least since the days of Chadwick and his followers, been considered systematically. Considerations of economic efficiency, of structural reliability and flexibility, and of uncertainties of costs and benefits were weighed in a kind of rudimentary balance. During the past century systems designs, based on intuitive judgments, have been gradually tempered by errors and successes with the evolution of "rules of thumb" to more publicly verifiable practices.

The decisions of engineers and planners in collecting and evaluating data and formulating rules for their use in design have been made in the context of objectives that often have been conflicting, poorly defined, and imperfectly connected with the external economy. Some of the design methodology of earlier eras contains assumptions now no longer tenable under close scrutiny, however expedient they were in their day. Many of these rules have stood the test of time and more searching analysis and are valid today despite shifts in public objectives and private desires. The recent, rapid development of the logic of planning and advances in computer capabilities have encouraged a new generation of engineers to aspire to create methods of systems analysis of ever increasing scope.

The members of the Harvard Water Resources group who worked on the project reported herewith included staff members Harold A. Thomas, Jr., Robert P. Burden, Myron B Fiering, Zane Spiegel, and Mrs. Nina Shahrik, and graduate students Walter O. Spofford, Jr., Joseph J. Harrington, Anton J. Muhich, David A. Robertson, Jr., and Peter Rogers. Maynard Hufschmidt, Director of Research of the Harvard Water Program in the Graduate School of Public Administration, and John M. Wilkinson, Research Associate in the Graduate School of Public Administration, together with

Dr. Fiering were responsible for the simulation model of the Lehigh River. Appreciative acknowledgement is made of the cooperation and participation of Dr. Richard L. Woodward, Project Officer on the contract for the Division of Water Supply and Pollution Control.

The work described in this report is a part of a broader coordinated research program, carried on by the Division of Engineering and Applied Physics and the Graduate School of Public Administration, with financial support also from the Corps of Engineers, United States Army, and Resources for the Future, Inc. Thus, for example, the simulation problem of Chapter 2 of this report was developed jointly for use in connection with the Corps of Engineers and Public Health Service research programs.

LIST OF FIGURES

	Figure Number	Title
Chapter 1.	1	Storage-Yield Relation Obtained by Sequent Peak Procedure
	2	Required Storage vs. Length of Record
	3	Required Storage vs. Length of Record
	4	Average Required Storage vs. Length of Record
	5	Distribution of Required Storage
Chapter 2.	1	The Lehigh River System
	2	Generalized Diagram of the Overall Simulation Program, Lehigh River
	3	Flow Chart - Monthly Operating Program
	4	Flow Chart - Subroutine BASIC
	5	Flow Chart - Subroutine POWER
	6	Flow Chart - Subroutine DIVERT
	7	Flow Chart - Subroutine WATER
	8	Flow Chart - Subroutine SUPPOW
	9	Flow Chart - Subroutine SPACE
Chapter 3.	1	Plan View of River Model
	2	Transition Matrix for Hydrology at Reach A
	3	Transition Matrix for Pollutant at Reach A
	4	Assumed Storage-Yield Relationship for Hydrology
	5	Conditional Probability Matrix of Flow and Storage in Reach A

List of Figures (cont'd)

	Figure Number	Title
Chapter 3.	6	Joint Probability Matrix of Flow into Reach B
	7	Transition Matrix of Flow into Reach B
	8	Storage Yield Relationship for Pollutant
	9	Amount of Pollutant Remaining after Decay, Given Amount before Decay
	10	Transition Matrix for Pollutant Assuming Pollutant Independent of Hydrology
	11	Transition Matrix for Pollutant with Dependence on Hydrology
	12	Simplified Notation for Conditional Probability Matrix of Flow and Storage Shown on Fig. 5
	A-1	Standard Notation
Chapter 4.	1	Relationship of Cost to Soil Capacity
	2	Relationship of Design Renewal Period to Design Soil Capacity
	3	Relationship of "Optimal" Cost to Design Soil Capacity for Different Values of "True" Soil Capacity
	4	Relationship of "Regret" to True Soil Capacity for Different Values of Design Soil Capacity
	5	Schematic County-Subdivision Layout

List of Figures (cont'd)

	Figure Number	Title
Chapter 4.	6	Typical log-Normal Plot of y-Statistics
	7	Relationship of Risk (per Installation) to Design Variate
	8	Design Variate vs. Coefficient of Variation for Different Values of β
Chapter 5.	1	Population Function
	2	Utility Function
	3	Cost Function
	4	Design Capacity - Net Benefit
	5	Design Capacity - Net Utility
	6	Net Benefits - Utility
	7	Design Capacity - Interest Rate
	8	Design Capacity - Economic Time Horizon
	9	Graphic Solution of Equation 3
Chapter 6.	1	Two-Dimensional Representation of the Three-Dimensional Surface
	2	Paratroopers Walking up the Constrained Surface
	3	Shape of the Surface for the Didactic Model
	4	Contours of Constrained Surface of Didactic Model
	5	Generalized Block Diagram, Constrained Steepest Ascent
	6	Subroutine CHECK, Didactic Model
	7	Behavior of Ascent near Constrained Maxima

LIST OF TABLES

	Table Number	Title
Chapter 1	1	Storage-Yield Relation: Deterministic Case
	2	Average and Standard Deviation of Required Storage
	3	Comparison of Theoretical and Measured Storage
Chapter 5	1	Conversion of Interest Rates
	2	Average Domestic Water Rates
	3	Determination of N for Various Values of Q_D
	4	Computation of Q_D for Various Values of g
	5	Computation of N for Various Values of g
	6	Computation of Design Capacity for Various Values of Interest Rate
	7	Computation of Design Capacity for Various Values of T
	8	Computation of y

1-a

OPERATIONS RESEARCH IN WATER QUALITY MANAGEMENT

INTRODUCTION

New methods are reported for evaluating the composite effects of combinations of streamflow regulation, waste water treatment, and waste water regulation for application in comprehensive programs to improve and maintain the quality of water in major water resources systems. The objective is to show how these methods may be used to strengthen the rational base for river basin planning by use of mathematical systems analysis. The work on problems of water quality management which will be described was part of a larger study on water resource planning carried out cooperatively by the Division of Engineering and Applied Physics and the Graduate School of Public Administration. The effort of the latter group was directed to aspects of water resource development other than water quality management. The joint program was centered on an analysis of the Delaware River system with particular concentration on the Lehigh River Basin.

Several new mathematical techniques useful in fusing engineering design, economic analysis, and governmental planning have recently been developed to exploit the potentiality of the electronic computer. The new mathematical constructs derive from game theory, statistical decision theory, linear programming, and dynamic programming. From the viewpoint of the administrator or engineer concerned with water quality management systems,the greatest utility of the new methodology stems from the fact that many more alternative designs covering a broad range of technologically feasible possibilities may be analysed and compared on a sound basis than could be done before the advent of the computer.

In this new area of engineering analysis there are many new ideas, tools, and techniques. In introducing them, it has seemed best to discuss them singly or in small groups in connection with several relatively simple applications rather than to deploy all of them simultaneously on a large and complex water quality management system. Accordingly this report is arranged in seven chapters in which the new techniques are exhibited and discussed individually. It should be stated, however, that while this arrangement is useful in attaining clarity of exposition it does not display the full thrust and versatility of the techniques of systems analysis when they are applied simultaneously to an intricate management system for water quality control. The material presented represents the first steps of infusion of operations research methodology in water quality management. Work on more sophisticated models and comprehensive management problems is proceeding at Harvard with the support of a Public Health Service Research Grant.

The methods and results of the seven chapters are briefly discussed in the following sections of this introduction. For ease of reference, a glossary is appended for readers who may not be familiar with some of the specialized technical terms used in operations research.

(1) Statistical Analysis of Reservoir Storage-Yield Relations

Dams are the chief devices for controlling that segment of the hydrosphere with which man is most directly concerned. The art of dam construction is in an advanced state, but the same cannot be said of the art of predicting the efficacy of control of the affected stretch of the rivers. A better understanding of the storage-yield relation for reservoirs is a basic need for the improved design of water resource systems. Heretofore, the technology of flow regulation has been based on hard-won experience and on reasoning from inappropriate and over-simplified probability models. The results have only a limited transfer value from one region to another because of the large number of variables and objectives involved in flow regulation.

The investigation described in this chapter is an important part of our development of synthetic hydrology methods for use in long-range water resource planning. The objective is to apply efficient statistical techniques to hydro-meteorological data to make available the maximum amount of information for planning and operation. Results are presented of an extensive computer study indicating the level of reliability of estimates of storage requirement based on mass-diagram (Rippl method) calculations of the type commonly made. In the analysis, a wide variety of possible statistical populations of streamflow were investigated together with a broad range of annual draft levels arising from water-uses such as municipal water supply, low flow regulation, recreation, power, navigation, and flood control. The reliability of conventional storage estimates made from streamflow records of the usual length is shown to be surprisingly low. An alternative method is described for determining storage requirements which extracts more efficiently the information inhering in flow records and which therefore, provides more precise estimates of storage requirements.

(2) Computer Simulation of Multipurpose Water Resource Systems

A mathematical model for computer simulation of the Lehigh River system is described which has been developed jointly by our project and the cooperating group in the Graduate School of Public Administration. This model takes into consideration streamflow variations over long time periods with concomitant changing demands for dilution water and for water supply with increasing population density and industrial development. The simulation model also takes into consideration other uses of water, including hydropower and recreation. Our program is a general one which can be used for development schemes in other river basins. The purpose of the model is to aid in the identification of that particular combination of flow-regulating structures, treatment plants, and other hydraulic works, which most nearly achieves the objectives of the comprehensive plan.

(3) Queueing Model for Pollution Transport in Streams

A preliminary stochastic model of the Markov type has been constructed to investigate problems associated with stream pollution. The objective is to predict, for any given set of reaches of a stream, the quality of water at specified downstream control points that may be expected with: Any pattern of stream hydrology; any pattern of pollutional loading; and any configuration of dams, tributaries, and water intakes. It is intended that this model, in its final form, will incorporate all the important processes of self-purification. In its present stage of development this model is able to trace pollutants having time-dependent decay, but as yet it does not embody sedimentation, absorption, biological uptake, photosynthesis, or bottom deposits. When completed it will be incorporated in the computer program (2) to improve simulation of intricate water resource systems.

(4) Application of Decision Theory to Sanitary Engineering Design Problems

A mathematical model employing statistical decision theory has been developed to aid in the planning of complex systems of sanitary engineering works that are to be installed in stages over a period of several years. The model employs matrix algebra and response surface sampling, and features a newly perfected method for reaching decisions in the face of physical uncertainties, errors in data, and economic risks. In order to show the utility of the model, it is applied to the problem of disposal of sewage from individual households in unsewered subdivisions and groups of subdivisions. The objectives are as follows: To develop a formulation and computational routine for determining the optimal balance (size and arrangement) of components - septic tank and absorption field devices - at each site so as to maximize economic efficiency; to develop criteria for establishing an ideal administrative code indicating the proper number and kind of tests to be made in new subdivisions and indicating the proper interpretation of these as regards the capacity of the absorption field; and to indicate in each application whether or not additional data on soil characteristics are needed and are economically justifiable. The model does this by a systematic evaluation of the costs of making incorrect decisions and of the costs of having only uncertain information - uncertainty such as that inhering in the measurement and interpretation of fresh water percolation tests and in the estimation of costs of land and sewage handling devices. The model takes into consideration economic parameters such as discount rates and time horizons; cost of financing and administration are included.

(5) Mathematical Model for Determining Optimal Sizes of Water Treatment Plants

In the construction of sanitary engineering works, such as water and sewage treatment plants, the problem arises of how much excess capacity to build at the present time to accomodate loads or demands from anticipated future population growth and industrialization. The problem is essentially the balancing of economies of scale in present construction against the diseconomy of unprofitable investment in excess capacity that will not be used for several years. A model using decision theory is presented which indicates the proper size of water plants. The plant size that maximizes economic efficiency is formulated as a function of population growth rate, the per capita demand for water, the interest rate for capital investment, and the economic time horizon. A new feature in this model is the introduction of a loss function which assigns economic penalties to water shortage of various magnitudes. It is basically unrealistic to treat the municipal water requirement as a rigid constraint since a large part of the water is used for lawn irrigation, recreation, and industrial purposes. When shortages occur these activities are curtailed; such curtailment does not necessarily increase health hazards. Therefore it is reasonable to predicate a loss function for short term economic losses incurred with small or medium-sized water shortages. The introduction of the loss function provides an element of flexibility in the design and removes the treatment of demand as a rigid constraint.

(6) A Method of Optimizing Designs of Waste Management Systems

In systems analysis of problems of water quality management, complex mathematical formulations of many variables are developed to reproduce or simulate prototype systems. In solving these, an attempt is made to identify that particular set of the decision variables (tank sizes, reservoir and aqueduct capacities, chemical dose rates, minimum flow releases, sampling frequency, etc.) that most nearly accords with the preset objectives of the plan and at the same time conforms to all relevant constraints (physical, chemical, institutional, and budgetary). Often formidable mathematical difficulties are encountered particularly when many variables and constraints must be dealt with simultaneously. A generalized technique for solving problems of this type has been developed and adapted to machine computation. In applying this technique, which is called the "method of steepest ascent", the computer evaluates the system at an arbitrary starting point corresponding to a particular set of assigned values of the decision variables that represent a physically possible design, but not necessarily one that is satisfactory from an economic point of view. Next the computer evaluates the gain or loss of economic efficiency that results from a small increment applied to each of the decision variables in turn. Having done this, it automatically computes that combination of increments (or decrements) to the variables that causes the net benefit to ascend most rapidly. In this new position the variables are incremented again and another step is made which increases economic efficiency still further. The process is repeated by "steepest ascent" movement on the "response surface" until the peak corresponding to the optimal design is attained. At this point a change in any of the design variables will result in a loss of economic efficiency. In complex systems there are often several "peaks", and in order to identify the best design is is necessary to repeat the foregoing process from several different starting points. In selecting these, use is made of a random sampling technique and a theorem of probability theory that indicates the number of steepest ascent runs to be made. In this chapter, the method is applied to a sample problem of design of a settling tank in order to show the details of the procedure. In practice, however, it may be used to solve much more sophisticated and difficult problems having dozens of variables and many constraints.

(7) Imputed Benefits of Municipal Water Supply

The difficulties of attaching explicit monetary benefits to municipal water supply are well known. In practical design the usual recourse is to treat urban water demand as a constraint. Water is routed to this use at the projected level without consideration of the demands associated with other uses. Commonly, the requirement for municipal use is based on projections of population growth and per capita demand. Little or no consideration is given to the degree of complementarity with other water uses or to demand elasticity which may be rather high for some components of municipal water use such as irrigation and recreation.

Until the time when valid externally determined benefit functions for varying levels of municipal water supply can be identified, some insight may be had into the problem of rational design of multipurpose systems involving urban water supply by consideration of the "imputed" benefit of this purpose which is implicit in the setting of a requirement as a constraint. The imputed marginal benefit is equal to the increment in benefits that would be added to other purposes having monetary benefit functions, if the urban water requirement were reduced one unit.

A simple water resource system is discussed in this chapter which includes two uses: water-power with monetary benefits; and municipal water supply without monetary benefits, but having a fixed water requirement treated as a constraint. The system is analyzed with each use considered separately in a single-purpose design, and then with both in a multipurpose design. The method of computation of both marginal and average imputed benefits is described and the use of these parameters as planning aids is briefly discussed. In an analogous manner imputed benefits may be calculated and used for certain other system purposes such as low flow regulation for quality control.

It is hoped that the results summarized in this report and those that we hope to achieve in the near future will serve to show the power of the methods of systems analysis as a timely innovation for improvement of the art of decision making in the field of water quality management.

CHAPTER ONE

Statistical Analysis of the Reservoir Storage-Yield Relation

The most important structures for the regulation of surface runoff are dams. Although man has had experience with these structures for three or four millennia in mitigating floods and droughts, and although he has developed ingenious and economical methods of construction for many types of sites, his methods of estimating the proper size dam for a given stream and target pattern of draft are based on "rules of thumb" that are at the present time more of an art than a science.

The common method for analyzing the relationship between reservoir inflow, desired draft, and storage capacity is that of W. Rippl(1). This procedure appears in all standard textbooks on water supply. Rippl's method, which is based on the mass diagram, assumes that both inflows and outflows are known functions of time and gives the minimum effective storage required such that no water shortage will occur during the time period under consideration. There are three defects in this method: the analysis is necessarily based on the historical record, and it is unlikely that the same flow sequence will occur again during the active life of the completed structure; the length of the historical record is apt to be quite different from the economic life of the dam, and since the storage capacity obtained by the Rippl method increases with the length of the record, the estimated capacity is likely to be incompatible with a design based on the economic life, which is determined by social and financial considerations rather than by physical considerations; and the mass-diagram method does not provide help to the designer in the setting of the calculated risk to be taken with regard to water shortages during periods of low streamflow.

---

[1] Rippl, W., "The Capacity of Storage Reservoirs for Water Supply," Proc. Institution of Civil Engineers 71, 270, 1883.

Several hydraulic engineers following Hazen(2) have presented methods for coping with this third defect. These methods are generalizations based on arithmetical operations studies with sets of standardized streamflow records. Different reservoir storage capacities are tested and counts are made of the number of periods of water shortage occurring with each size. Hazen's tables give a useful first approximation to the desired reservoir size. They require only knowledge of the coefficient of variation of streamflow (standard deviation of annual flow divided by the mean annual flow) and of the projected level of development of the stream (annual target draft divided by mean annual flow). However the application of Hazen's tables and the tables and nomograms of others who have elaborated on this approach is limited in the following respects: (1) only a relatively small fraction of the possible types of stream inflow hydrology have been studied - graphs useful for one region are not necessarily applicable in another; (2) the results are based on the assumption of a uniform draft rate throughout the year, whereas for some purposes such as irrigation the annual pattern of draft varies markedly from season to season; and (3) the frequency of water shortages is not precisely estimated owing to the shortness of most existing streamflow records. With fifty years or less of historical stream traces, it may well happen that the most severe floods and droughts do not provide an adequate indication of the magnitude of floods and droughts that in fact will occur during the economic life of the dam.

A method of contending with this limitation is to use probability theory and statistical analysis. This method assumes that the existing record of streamflows is only one sample drawn from an infinite set of possible samples that could be generated from a population of streamflows. Each stream has its own population; adjacent streams have populations that resemble each other in

---

[2] Hazen, Allen, "Storage to be Provided in Impounding Reservoirs for Municipal Water Supply," Trans. Am. Soc. Civil Engrs. 77, 1539 (1914).

some respects but are not identical. Sudler(3), Barnes(4) and others have employed decks of cards and other sampling devices for generating synthetic flow traces. On each card is printed a "representative" annual flow. By dealing from such a deck it is possible to create a long artificial record which may be subdivided into shorter "records" of a duration equal to the economic life of the prototype dam. These may be analyzed separately by the mass diagram method and a probability distribution of reservoir capacities obtained. From this a design value may be selected such that, for example, the reservoir may be expected to run dry only once in twenty years on the average. The central idea of the mathematical synthesis is to create the critical periods of high and low runoff that are probably not subsumed in the short historical records, but which, from the viewpoint of probability theory, could be expected to occur in an actual record of sufficient length.

In the work of our Water Resources Seminar we have developed and extended the method of synthetic hydrology (Thomas and Fiering(5), Fiering(6)). Our method of synthesizing streamflow sequences, which is designed for electronic digital computation, has the following advantages over previous methods: it may be used for hourly, daily, weekly, monthly or seasonal flows as well as for annual flows; it does not require that the probability distribution for each time period be a normal distribution but rather permits of a reproduction of

---

[3] Sudler, Charles E., "Storage Required for the Regulation of Stream Flow," Trans. Am. Soc. Civil Engrs. 91, 622 (1927).

[4] Barnes, F. B., "Storage Required for a City Water Supply," J. Inst. Engrs., Australia 26, 198 (1954).

[5] Thomas, H. A., Jr., and M. Fiering, "Synthetic Streamflow Sequences" Chapter 12, Design of Water Resource Systems. Harvard University Press, 1962.

[6] Fiering, M. B, "Queueing Theory and Simulation in Reservoir Design," Transactions of the American Society of Civil Engineers, V. 127, Part 1, No. 3367, 1962.

the skewness and kurtosis of the empirical distributions as calculated from the record; it incorporates serial correlation between flows in successive time periods to accord with that manifest in the record; and finally by means of a cross-correlation technique the method is adapted to the design of multi-reservoir projects.

A mathematical description of the stochastic process that we advocate for the generation of synthetic flow traces is given in a subsequent section of this chapter. At this point, however, it is pertinent to try to answer two questions that arise. These are the following: (i) What is the justification for the assumption that a population of streamflows exists and that the probability method is valid? (ii) Even if probability theory is applicable to hydrological sequences, does it follow that the stochastic method of synthetic hydrology for setting reservoir capacities leads to better designs than the deterministic method of the mass-diagram?

Our answer to the first question is that extensive statistical analyses of historical traces of many streams in all parts of the world show patterns that are characteristic of samples generated synthetically by stochastic processes. For example, if a fifty year record of annual flows at a gaging station on a stream with a reasonably stable hydrologic regime is subdivided into ten successive subrecords of five years each which are analyzed individually, it may be expected that the ten means of the subrecord will be distributed in consonance with the theoretical distribution of such means as derived from an appropriate stochastic process. Likewise probability functions for the standard deviation, the extreme values, and the skewness can be derived and they will accommodate the corresponding parameters of the subrecords within expected allowance for statistical fluctuation.

We make two reservations, however, as to the potentialities of synthetic hydrology. In the first place the method cannot be used reliably on streams of unstable regime, or those having watersheds appreciably altered by activities of man during the period of record. In the second place by "an appropriate stochastic process" we do not mean static probability functions or random frequency distributions such as have been used to formulate flood frequency curves or duration series. To model nature, mathematical formulation of hydrological time series requires the incorporation of serial correlation between successive events. There are several ways that this may be done. For the time periods of ten to a hundred years used in water resource planning, the simple autoregressive series that we have used and which will be described appears to be useful for most design purposes. For larger periods of time such as those studied by Hurst(7) this simple formulation is not valid and more elaborate models are needed. Hurst's analysis of extensive records of natural phenomena including rainfall, river statistics, tree rings and mud varves, indicated long-term shifts in the mean and other non-random effects that cannot be accounted for by models with sets of simple correlations. These shifts are undoubtedly the effect of large scale trends and fluctuations in global climatological patterns. While the development of a sophisticated mathematical model to explain these complex series of figures would be of considerable interest in theoretical hydrology and geophysics, it is of less importance to the problem of economic planning of water management systems.

The answer to the second question as to why, or in what sense, synthetic hydrology gives a more valid basis for design than the conventional tools of the mass-diagram and short-term operations studies with historical flows is that by using the methods of synthetic hydrology it is possible to extract more efficiently the relevant information that inheres in the record. The

---

[7] Hurst, H. E., "Methods of Using Long-Term Storage in Reservoirs," Proc. Institution of Civil Engineers, Part I, vol. 5, 519-590, Sept. 1956.

results of the computer investigation to be described shows that the mass-diagram method uses the observed data inefficiently and yields results that are statistically unstable. The same is true of drought-frequency estimates based on short-term operations studies of historical flow sequences, but this will not be demonstrated in the present report.

In the following paragraphs we shall present an exact mathematical treatment of the problem commonly solved by the mass-diagram, discuss some pertinent theoretical results obtained by Feller, describe the computer program, and discuss the results.

The Deterministic Solution of the Storage-Yield Relation by the Mass-Diagram.

The problem is the following: Given a record of streamflows at the site of a proposed dam for a period of T years, and also given the desired drafts for this period, it is required to find the minimum capacity of reservoir such that no shortage in draft will occur. The following solution assumed that the inflow and drafts will be repeated in a cyclic progression in successive sets of cycles of T-years each. In the solution only two cycles are needed.

Let W_i be the amount wasted in the i^{th} inverval, X_i be the i^{th} inflow, and D_i be the i^{th} draft. The total volume to be wasted in any cycle is

$$\sum_{i=1}^{T} W_i = \sum_{i=1}^{T} X_i - \sum_{i=1}^{T} D_i \qquad (1)$$

The solution of the problem entails the identification of the optimal program of wasting.

Procedure

1. Calculate $X_i - D_i$ for each time period $1 \leq i \leq 2T$; and calculate the net inflows, $\sum_{i=1}^{t} (X_i - D_i)$, $1 \leq t \leq 2T$ (see column 7 of the computation of Table 1).

2. Locate the first peak, P_1, (maximum) in the column of cumulative net inflows (column 7).

3. Locate the "sequent" peak, P_2, which is defined as the next peak having a larger magnitude than the first.

4. Between this pair of peaks find the lowest trough (minimum), T_1, and calculate, $P_1 - T_1$.

5. Starting with P_2, find the next sequent peak, P_3, which is the next peak that has a greater magnitude than P_2.

6. Find the lowest trough, T_2, between P_2 and P_3 and calculate $P_2 - T_2$.

7. Starting with P_3 find P_4 and T_3; calculate $P_3 - T_3$.

8. Repeat with P_4, etc. for all sequent peaks in the series for 2T years.

9. The required reservoir size is

$$S_m = \text{Max } (P_j - T_j) = P_m - T_m, \qquad (2)$$

where j is the index for sequent peaks and troughs.

10. The reservoir will be empty at the end of the time period of the occurrence of the minimum trough, T_m, between the pair of sequent peaks, P_m and P_{m+1}. At the end of the next time period the storage will be increased by an amount equal to the net inflow provided that this net inflow is less than S_m. If it is greater than S_m, the storage will be increased to S_m and the excess wasted.

11. The storage, S_i, at the end of the i^{th} time period (season) and the waste flow during this season may be calculated from the following continuity relations:

$$S_i = \text{Min } [S_m, (S_{i-1} + X_i - D_i)] \qquad (3)$$

$$W_i = \text{Max } [0, (X_i - D_i - S_{i-1})] \qquad (4)$$

The computation is exhibited in the last three columns of the computation. During each cycle the reservoir will be empty at least once and be full at least once.

Example.

The method is illustrated for a 12-month record of flows to show details of the computation. In actual use and in the computer program it is of course applied to much longer periods of time. The draft is set at an average rate of 90% of the average rate of inflow. The required storage is 14.8×10^3 acre-feet.

On Figure 1 this result and those other similar computations at various draft rates are plotted to give the storage-yield formulation for the stream for this particular draft pattern. In the computer program this algorithm for determining the storage requirement is called the "sequent peak" procedure.

Theoretical Considerations

The treatment of the storage-yield relation by the methods of mathematical statistics is beset with difficulties and only a few significant results are available at the present time to aid the engineer in the design of water resource systems. Feller(8) in 1951 derived results for the asymptotic distribution of the range of the cumulative departure of a stochastic variate from its mean. In terms of the storage-yield relation his results apply to cases where the variate is the annual flow of a stream, a hundred percent development is prescribed, (annual draft equal to mean annual flow) and the time period is long. Feller's theory does not indicate how long the period must be such that his results are valid, but our computer runs indicate that a period of fifty years is sufficient for most types of streams. Anis and Lloyd(9)

---

[8] Feller, William, "The Asymptotic Distribution of the Range of Sums of Independent Random Variables," Annals of Math. Stat. 22, 427-432, Sept. 1951.

[9] Anis, A. A., and E. H. Lloyd, "On the Range of Partial Sums of a Finite Number of Independent Normal Variates," Biometrika, 40, 35-47, 1953.

developed an approximate formulation for the moments of the range for small sequences, but no solutions have attempted to account for the effects of non-normal flows, serial correlation, and regulation less than 100%.

The range of the cumulative departure of the mean is defined as follows:

Let $z_t = \sum_{i=1}^{t} (x_i - \bar{x}), \quad i = 1, 2, \ldots T; \quad 1 \leq t \leq T$

where x_i = the annual flow of stream during year i

$\bar{x}$ = the mean annual flow during T years, also the mean draft rate

z_t = cumulative departure from the mean at the end of year t.

Then the range of the cumulative departure from the mean, which is the indicated overyear storage requirement, is

$$R = \text{Max} \left| z_j - z_k \right| \qquad (5)$$

where $j = 1, 2, \ldots, T$

$k = 1, 2, \ldots, T$

The quantity R is a stochastic variate since it is computed from the x_i which are stochastic variates. Feller gives the first two moments of the asymptotic (large T) distribution of R which are valid for x_i that come from continuous probability functions having finite moments. The mean and standard deviation of R are as follows:

$$E(R) = (\pi/2)^{1/2} \sigma T^{1/2} = 1.2533 \sigma T^{1/2} \qquad (6)$$

$$SD(R) = (\pi^2/6 - \pi/2)^{1/2} \sigma T^{1/2} = 0.2723 \sigma T^{1/2} \qquad (7)$$

The coefficient of variation is

$$CV(R) = (\pi/3-1)^{1/2} = 0.217 \qquad (21.7\%) \qquad (8)$$

These theoretical results have significant implications if (as our computer study shows) the time period T does not have to be more than fifty or a hundred years for the asymptotic formulation to apply. They show that the range depends only upon one stream population parameter, namely the standard deviation of annual flows. Other parameters such as skewness have no importance. Also they show that the storage requirement increases in proportion to the square root of the time period. Therefore if the design is to be based on the logic of mass-diagram analysis (which we disparage) then the problem of the designer is (1) to get the most efficient estimate of the standard deviation by an appropriate statistical analysis of the record, and (2) to correct the result based on the time period of the record to the economic life of the structure using the square root law. Fiering(10) has presented an efficient method for the estimation of the standard deviation which employs a correlation technique and uses information from adjacent gaging stations. Neither of these requires the construction of mass-diagrams.

Another important implication of Feller's results is that the range is quite unstable statistically. For example, if the expected value of the range E(R) were 100 and we assume that the distribution of R is approximately normal, then there is about one chance in six that the historical record will yield a result lying outside the range 100 ± 22 (less than 78 or more than 122); there is about one chance in forty that the result will fall outside the range 100 ± 2 x (22) (less than 56 or more than 144). It is evident therefore that the value of the range obtained from an observed record (whether by the mass-design or by the more efficient procedure discussed) is apt to be erratic owing to non-significant idiosyncrasies of that record. However before one can reach a full evaluation of the mass-diagram technique it is necessary to consider some aspects not covered by the foregoing theory. The following questions may be asked. For short time periods what is the effect of skewness and serial correlation of the streamflow population on the storage requirement? For designs with low levels

---

[10] Fiering, M. B, "Use of Correlation to Improve Estimates of the Mean and Variance," U. S. Geological Survey, Professional Paper, 434-C, 1963.

of development for which overyear storage is unimportant, what streamflow parameters are significant? Is the square law valid for short-time periods? How stable statistically are estimates of R based on short-time periods? Our computer program was designed to answer these questions and others pertaining to the statistical stability and efficiency of estimates of drought frequencies based on short-term reservoir operations studies based on observed stream records.

The Hydrologic Model

At the start we postulate for stream flow a linear regression model of lag one; that is, the flow in any season is a linear function of the flow in the preceding season. Thus

$$Q_{i+1} = \bar{Q}_{j+1} + \beta_j (Q_i - \bar{Q}_j) \qquad (9)$$

wherein Q_{i+1} and Q_i are the discharges during the $(i+1)^{st}$ and i^{th} season, respectively, reckoned from the start of the synthesized sequence; $\bar{Q}_{j+1}$ and $\bar{Q}_j$ are the mean seasonal discharges during the $(j+1)^{st}$ and j^{th} season, respectively, within a repetitive annual cycle of seasons; and β_j is the least squares regression coefficient for estimating the flow in the $(j+1)^{st}$ from the j^{th} season.

The expected value of Q_{i+1} is $\bar{Q}_{j+1}$; thus the first moment of the distribution of historical flows is preserved by the model in Equation 5. However, if β_j is replaced by $\rho_j \sigma_{j+1}/\sigma_j$, it can be shown that the variance $V(Q_{i+1}) = \rho_j^2 \sigma_{j+1}^2$. In this formulation σ_j and σ_{j+1} are the standard deviations of discharges during the j^{th} and $(j+1)^{st}$ season, respectively, and ρ_j is the serial correlation coefficient between flows in the j^{th} and $(j+1)^{st}$ season. The model stipulated by Equation 9 does not preserve the variance of the historical flows unless $\rho_j = \pm 1$, and some modification is necessary.

If t_i is a random variate with zero mean and unit variance, the recursion relation

$$Q_{i+1} = \bar{Q}_{j+1} + \beta_j (Q_i - \bar{Q}_j) + t_i \sigma_{j+1} (1 - \rho_j^2)^{1/2} \qquad (10)$$

preserves the mean and variance of the distribution of historical flows. Equation 10 entails the deterministic, or trend line, estimate defined by the linear relation of Equation 9 and a stochastic, or random, component proportional to the standard error of estimate or $\sigma_{j+1}(1 - \rho_j^2)^{1/2}$. We can envision bands of width 1, 2,...,n $(\sigma_{j+1}(1 - \rho_j^2)^{1/2})$ which contain a corresponding proportion of all the points in the bivariate distribution of Q_j and Q_{j+1}; the evaluation of these proportions depends upon the distribution of seasonal flows.

If a normal distribution is postulated, then the t_i must be standard normal variates, and Equation 10 is a sufficient algorithm because the first two moments are preserved. If the flows are characterized by the gamma distribution, 3 parameters must be maintained. The distribution of t_i must reflect γ, the coefficient of skewness, inherent in the seasonal flows.

Rewriting Equation 10 and cubing both sides,

$$\left[\frac{Q_{i+1} - \bar{Q}_{j+1}}{\sigma_{j+1}}\right]^3 = \left[\rho_j \frac{(Q_i - \bar{Q}_j)}{\sigma_j} + t_i (1 - \rho_j^2)^{1/2}\right]^3 \qquad (11)$$

and taking expectations

$$\gamma_{j+1} = \rho_j^3 \gamma_j + (1 - \rho_j^2)^{3/2} \; E(t_i^3)$$

whereupon it follows that, in order to preserve the skewness of the distribution of historical flows, the skewness of the random variate t_i is

$$\gamma_t = \frac{\gamma_{j+1} - \rho_j^3 \gamma_j}{(1 - \rho_j^2)^{3/2}} \qquad (12)$$

If t_N is a normal variate with zero mean and unit variance, t_N can be transformed into a variate t_γ whose distribution is approximately like gamma.(11)

---

[11] Kenny, J. F. and E. S. Keeping, Mathematics of Statistics, Vol. II, pp. 99, 105-107, Van Nostrand (1951)

The transform is

$$t_{\gamma} = \frac{2}{\gamma_t}\left(1 + \frac{\gamma_t t_N}{6} - \frac{\gamma_t^{2}}{36}\right)^{3} - \frac{2}{\gamma_t} \qquad (13)$$

which is valid for positive or negative skewness. This is verified in the computer runs by the fact that the skewness of the inflow distribution is reproduced, regardless of sign.

It is pertinent to note that more skewness must be built into the variate t_{γ} than exists in the original record so as to preserve the natural skewness of the latter. If we consider a one-season model for simplicity, and replace subscripted variables by unsubscripted values, Equation 12 becomes

$$\gamma_t = \gamma\,(1 - \rho^3)\;/\;(1 - \rho^2)^{3/2} \qquad (14)$$

whereupon $\gamma_t/\gamma \geq 1$. This result can be generalized to a multi-season model.

If the historical record were characterized by a 4-or more-parameter distribution, the kurtosis and higher moments of the distribution could be preserved by adjusting the distribution of t_i. In these studies the skewness is the highest moment preserved by the synthesized data.

The general sequence of computation can be summarized as follows:

(i) analyze the historical data and compute the constants $\bar{Q}_j$, β_j, ρ_j, σ_j, γ_j for each season of the proposed model;

(ii) specify whether the normal or gamma distribution is relevant - that is, whether skewness is to be neglected or preserved;

(iii) synthesize a flow sequence of desired length, each element of which is computed from a trend-line estimate and a random component characterized by the relevant input parameters and distribution;

(iv) examine the moments of the synthesized data to ascertain that the model faithfully, but not necessarily exactly, reproduces the given moments.

The essential property of every generated sequence is that, by all statistical tests of significance, the sequence cannot be distinguished from the original historical record. If the synthetic generating scheme contained in Equation 10 is valid, any sequence generated therefrom could be a sample from the total population of all possible records which might be drawn from a set of stream flows characterized by the moments of the observed sequence. Moreover, the required storage computed from any synthetically generated sequence is just as likely to be the storage required during the lifetime of the project as that associated with any other generated sequence. This includes the storage computed from the historical record; that is, the historical record defines a storage which is one member of the population of all possible storages. If a large number of replicate sets of synthetic data are generated, say 100, the resulting 100 storages define a distribution which is an approximation to the distribution of all possible storages. When 100 sets are used, the first two moments of the required storage are fairly reliable; the third, fourth, and higher moments of the distribution are not well defined.

Let us summarize the strategy. Using any one of the 144 combinations of skewness, serial correlation, length of record, and degree of regulation, 100 replicate sets of synthetic stream flow data are generated. For each set we compute the storage required to meet the specified draft. This required storage is a member of a sample of size 100, taken from the population of all possible required storages. This sample is used to define the distribution (i.e., the moments) of the required storage corresponding to that particular combination of model parameters.

Computer Studies

We consider four streamflow distributions, four record lengths, three levels of annual serial correlation, and three degrees of regulation. Since the number of parameter combination increases in a multiplicative fashion, 144 combinations define the sample space for this investigation. In each instance the annual flows are assumed to be derived from a population with a mean of 1.0 and a standard deviation of 0.25; the resulting coefficient of variation, 25%, is typical for streams in several regions of the United States. The computer is coded to generate a variable number of replicate sets of synthetic data; for each combination we choose 100. Thus 14,400 sets of hydrology are generated and analyzed by the computer program within a few minutes. The parameters assigned in the computation are:

Populations

Normal

Gamma with skewness -0.5

Gamma with skewness 0.5

Gamma with skewness 1.0

Record Lengths

10, 25, 50, and 100 years

Annual Serial Correlations

0.0, 0.1 and 0.2

Degrees of Regulation

100%, 90% and 80%

A one-season model is utilized throughout and only high degrees of regulation are considered in the studies summarized in this report. Thus the results pertain to overyear storage. The investigation of multiseason cases with low as well as high degrees of regulation will be reported subsequently. The computer generates a synthetic flow sequence of given length, from which the mean annual inflow is computed. The annual draft rate is computed from the sample mean (i.e., mean annual inflow) whereupon the sequent peak procedure produces a required storage. The program computes and reports the mean and standard deviation, for the sample of size 100, of the required storage. These results are summarized in subsequent paragraphs and in Table 2.

The most important result is the verification of the theory of Hurst and Feller for the case of 100% development. Table 3 extracts typical results from Table 2 and arrays them against their theoretical counterparts. The agreement for small values of N is fairly close and becomes very much better as N increases. Computations based on the method of Anis and Lloyd are only slightly better for N = 10 than those of Hurst. In general, the values at the low end of the range of calculated average storage, Table 3, are derived from the runs made with no serial correlation; the upper values are derived from runs made with a serial correlation of 0.2. This is in agreement with engineering experience, and it is known from the results of queueing theory that serial correlation tends to increase the required storage.

N, yrs.	E(Storage) $1.25\,\sigma\sqrt{N}$	Range of Calculated Average Storage	Range of Calculated C_v, %
10	0.99*	0.7-0.7	33-36
25	1.56	1.2-1.5	28-32
50	2.22	1.9-2.2	23-27
100	3.13	2.9-3.3	23-27

* Values in this column are given by Equation 6

Comparison of Theoretical and
Measured Storage, 100%
Regulation, 100 sets

Table 3

Other pertinent results include the apparent stabilization of the variance for low values of regulation. The length of record has relatively little effect on the absolute value of the variance, although the coefficient of variation decreases sharply as the record length increases. For 90% regulation, the sensitivity of the coefficient of variation is intermediate between the two extremes.

The relative unimportance of skewness is clearly demonstrated. Hurst's work, based on the normal distribution, and Feller's asymptotic solution both depend on the variance alone. The results of this research indicate that the variance is by far the most important parameter even for relatively short record lengths. Figures 2-5 show typical results, including the distribution of required storage for computer runs consisting of 100 replicate sets.

Discussion

The principle implication of this study lies in the rejection of the use of the single historic trace for the determination of the storage-yield relationship at a particular site. The storage required to meet any given yield is better treated as a statistical variate, characterized by a distribution which has the same first three moments as the observed stream flows. Synthetic hydrology in tandem with the sequent-peak method of evaluating the storage-yield characteristics for a particular trace provide a method for defining the pertinent statistics of the distribution of required storage. Given this distribution, or its moments, the specification of storage at a particular site is a problem in decision-making under risk. Such problems can be handled only when pertinent economic information, in the form of utility or loss functions, is used to augment the statistical results.

The use of synthetic hydrology to refine the concept of the storage-yield relationship is the subject of continuing research at Harvard University. We expect to publish a comprehensive summary of this aspect of our work in the near future.

Example: Annual Draft = 0.90 Annual Inflow. Annual Inflow = 166 Kiloacre Feet

Yr. No.	Seas. No.	Inflow, X	%Annual Draft	Draft, D	X-D	Σ(X-D)	Storage, S	Waste, W
						100.000		
1	1	10	4	5.976*	4.024			0
						104.024	4.024	
	2	12	4	5.976	6.024			0
						110.048	10.048	
	3	21	5	7.470	13.530			8.794
						123.578 P_1	14.784	Full
	4	8	6	8.964	-0.964			0
						122.614	13.820	
	5	6	9	13.446	-7.446			0
						115.168 T_1	6.374	
	6	19	10	14.940	4.060			0
						119.228	10.434	
	7	22	15	22.410	-0.410			0
						118.818	10.024	
	8	29	11	16.434	12.566			7.806
						131.384 P_2	14.784	Full
	9	11	9	13.446	-2.446			0
						128.938	12.338	
	10	3	8	11.952	-8.952			0
						119.986	3.386	
	11	14	7	10.458	3.542			0
						123.528	6.928	
	12	11	12	17.928	-6.928			0
						116.600 T_2	00.000	Empty

Yr. No.	Seas. No.	Inflow, X	%Annual Draft	Draft, D	X-D	Σ(X-D)	Storage, S	Waste, W
2	1	10	4	5.976	4.024			
						120.624		
	2	12	4	5.976	6.024			
						126.648		
	3	21	5	7.470	13.530			
						140.178 P_3		Full
	4	8	6	8.964	-0.964			
						139.214		
	5	6	9	13.446	-7.446			
						131.768 T_3		
	6	19	10	14.940	4.060			
						135.828		
	7	22	15	22.410	-0.410			
						135.418		
	8	29	11	16.434	12.566			
						147.984 P_4		Full
	9	11	9	13.446	-2.446			
						145.538		
	10	3	8	11.952	-8.952			
						136.586		
	11	14	7	10.458	3.542			
						140.128		
	12	11 2(166)	12 2(100)	17.928	-6.928			
						133.200		Empty

* 0.04 (0.90) (166) = 0.04 (149.40) = 5.976

$$S_m = P_m - T_m = P_2 - T_2 = 131.384 - 116.6 = \underline{14.8\text{ kaf}}$$

required storage

Storage-Yield Relation: General Formulation for
Deterministic Case with Known Inflows and Drafts

Table 1

Results of Computer Studies

Population	ρ	N, yrs.	Regulation, %		
			100	90	80
Normal	0.0	10	0.7 (.23)*	0.4 (.17)	0.2 (.17)
		25	1.2 (.35)	0.6 (.22)	0.4 (.18)
		50	2.0 (.47)	0.8 (.28)	0.5 (.18)
		100	2.9 (.70)	1.0 (.25)	0.6 (.19)
	0.1	10	0.7 (.24)	0.4 (.18)	0.2 (.17)
		25	1.3 (.37)	0.6 (.25)	0.4 (.20)
		50	2.1 (.54)	0.9 (.33)	0.5 (.21)
		100	3.2 (.74)	1.1 (.36)	0.6 (.22)
	0.2	10	0.7 (.29)	0.4 (.20)	0.2 (.18)
		25	1.4 (.43)	0.7 (.29)	0.4 (.21)
		50	2.2 (.53)	1.0 (.39)	0.5 (.24)
		100	3.2 (.74)	1.3 (.44)	0.7 (.24)
$\Gamma(+\frac{1}{2})$	0.0	10	0.7 (.23)	0.3 (.16)	0.2 (.13)
		25	1.2 (.35)	0.5 (.20)	0.3 (.14)
		50	1.9 (.46)	0.7 (.25)	0.4 (.13)
		100	2.9 (.73)	0.9 (.23)	0.5 (.14)
	0.1	10	0.7 (.24)	0.3 (.17)	0.2 (.14)
		25	1.3 (.39)	0.6 (.23)	0.3 (.16)
		50	2.1 (.55)	0.8 (.29)	0.4 (.16)
		100	3.2 (.75)	1.0 (.32)	0.5 (.17)
	0.2	10	0.7 (.30)	0.3 (.18)	0.2 (.15)
		25	1.4 (.44)	0.6 (.26)	0.3 (.17)
		50	2.2 (.54)	0.9 (.35)	0.4 (.19)
		100	3.2 (.75)	1.2 (.40)	0.5 (.19)

Population	ρ	N, years.	Regulation, %		
			100	90	80
Γ (+1)	0.0	10	0.7 (.24)	0.3 (.17)	0.1 (.10)
		25	1.2 (.36)	0.5 (.19)	0.2 (.10)
		50	1.9 (.46)	0.6 (.22)	0.3 (.09)
		100	2.9 (.78)	0.8 (.20)	0.3 (.10)
	0.1	10	0.7 (.26)	0.3 (.18)	0.1 (.11)
		25	1.3 (.41)	0.5 (.22)	0.2 (.12)
		50	2.1 (.58)	0.7 (.26)	0.3 (.11)
		100	3.2 (.78)	0.9 (.30)	0.4 (.13)
	0.2	10	0.7 (.32)	0.3 (.19)	0.1 (.14)
		25	1.5 (.47)	0.5 (.24)	0.2 (.14)
		50	2.2 (.57)	0.8 (.31)	0.3 (.16)
		100	3.3 (.78)	1.0 (.37)	0.4 (.15)
Γ (- $\frac{1}{2}$)	0.0	10	0.7 (.24)	0.4 (.20)	0.2 (.17)
		25	1.2 (.36)	0.6 (.24)	0.4 (.23)
		50	2.0 (.49)	0.9 (.31)	0.6 (.24)
		100	2.9 (.68)	1.1 (.29)	0.7 (.24)
	0.1	10	0.7 (.27)	0.4 (.20)	0.2 (.17)
		25	1.3 (.37)	0.7 (.29)	0.5 (.25)
		50	2.1 (.52)	1.0 (.37)	0.6 (.27)
		100	3.2 (.74)	1.2 (.41)	0.8 (.27)
	0.2	10	0.7 (.29)	0.4 (.27)	0.2 (.22)
		25	1.4 (.44)	0.7 (.32)	0.4 (.26)
		50	2.2 (.56)	1.0 (.43)	0.7 (.29)
		100	3.2 (.72)	1.4 (.48)	0.8 (.29)

* Entries in the body of the table are the average storage based on a sample of 100; the associated standard deviation is in parenthesis.

Average and Standard Deviation of Required Storage,
100 Replicate Sets

Table 2

1-22

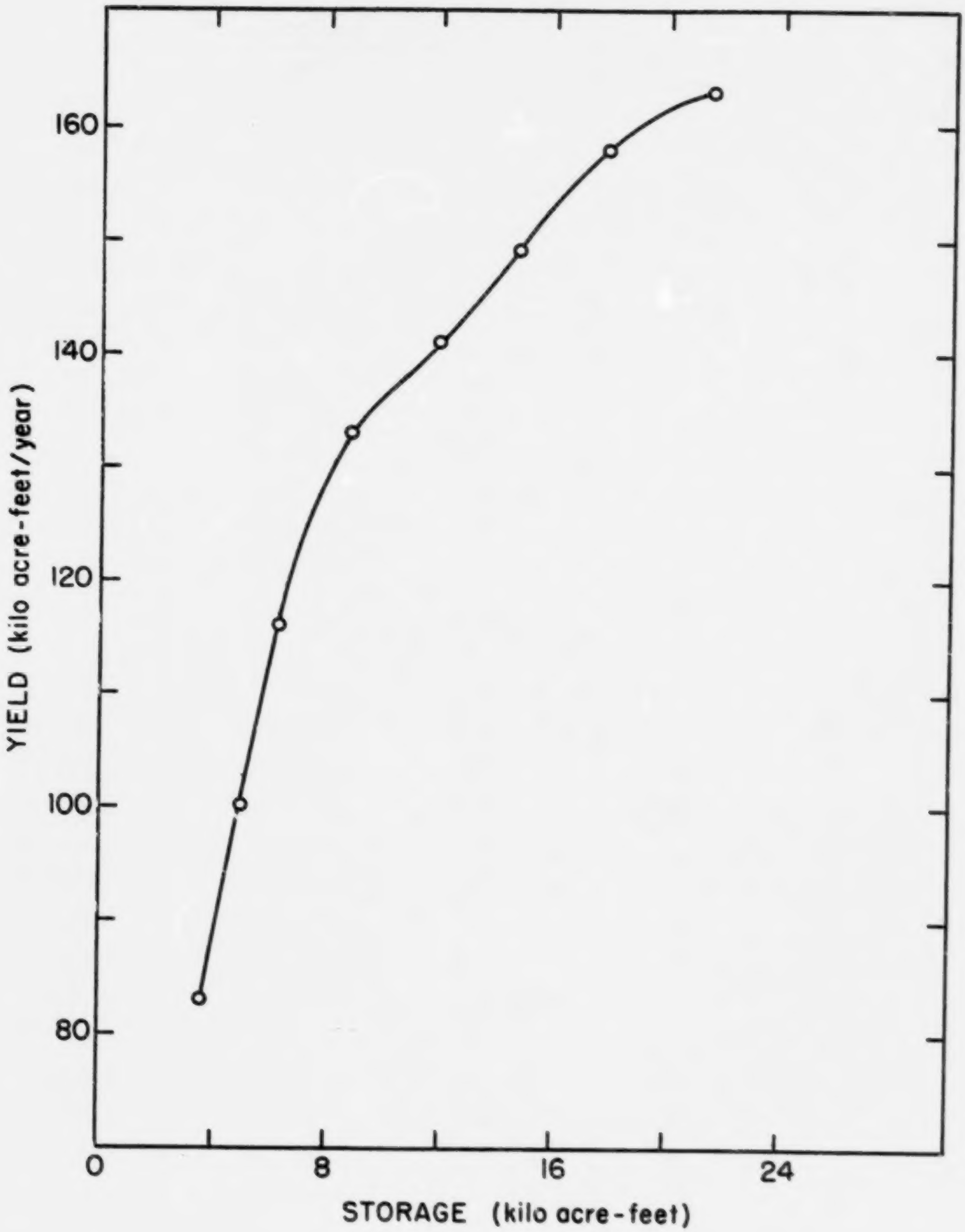


FIG. 1 STORAGE-YIELD RELATION OBTAINED BY THE SEQUENT PEAK PROCEDURE

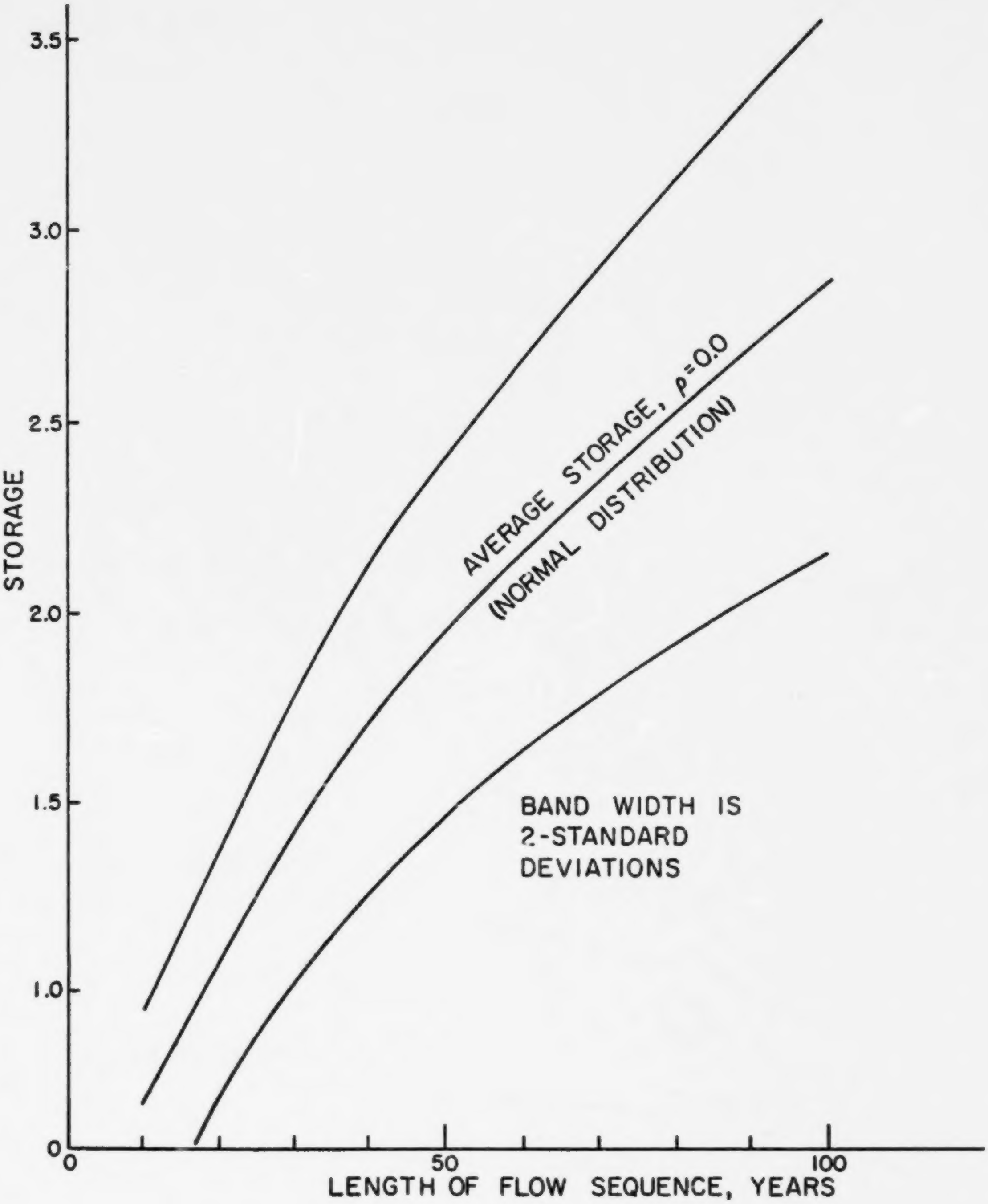


FIG. 2 REQUIRED STORAGE vs. LENGTH OF RECORD, 100% REGULATION

1-24

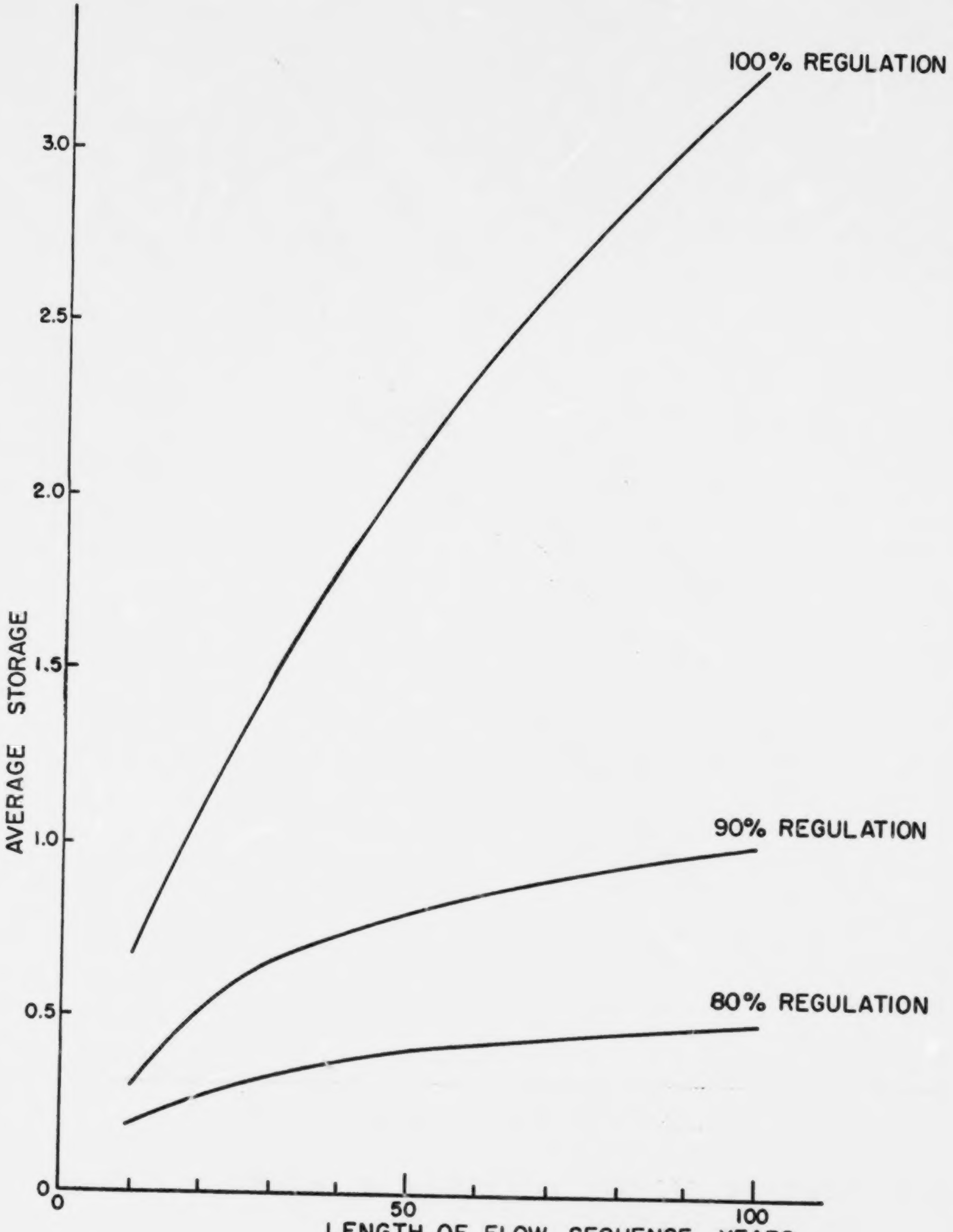


FIG. 3 AVERAGE REQUIRED STORAGE vs. LENGTH OF RECORD, $\rho=0.1$, GAMMA DISTRIBUTION ($\gamma=0.5$)

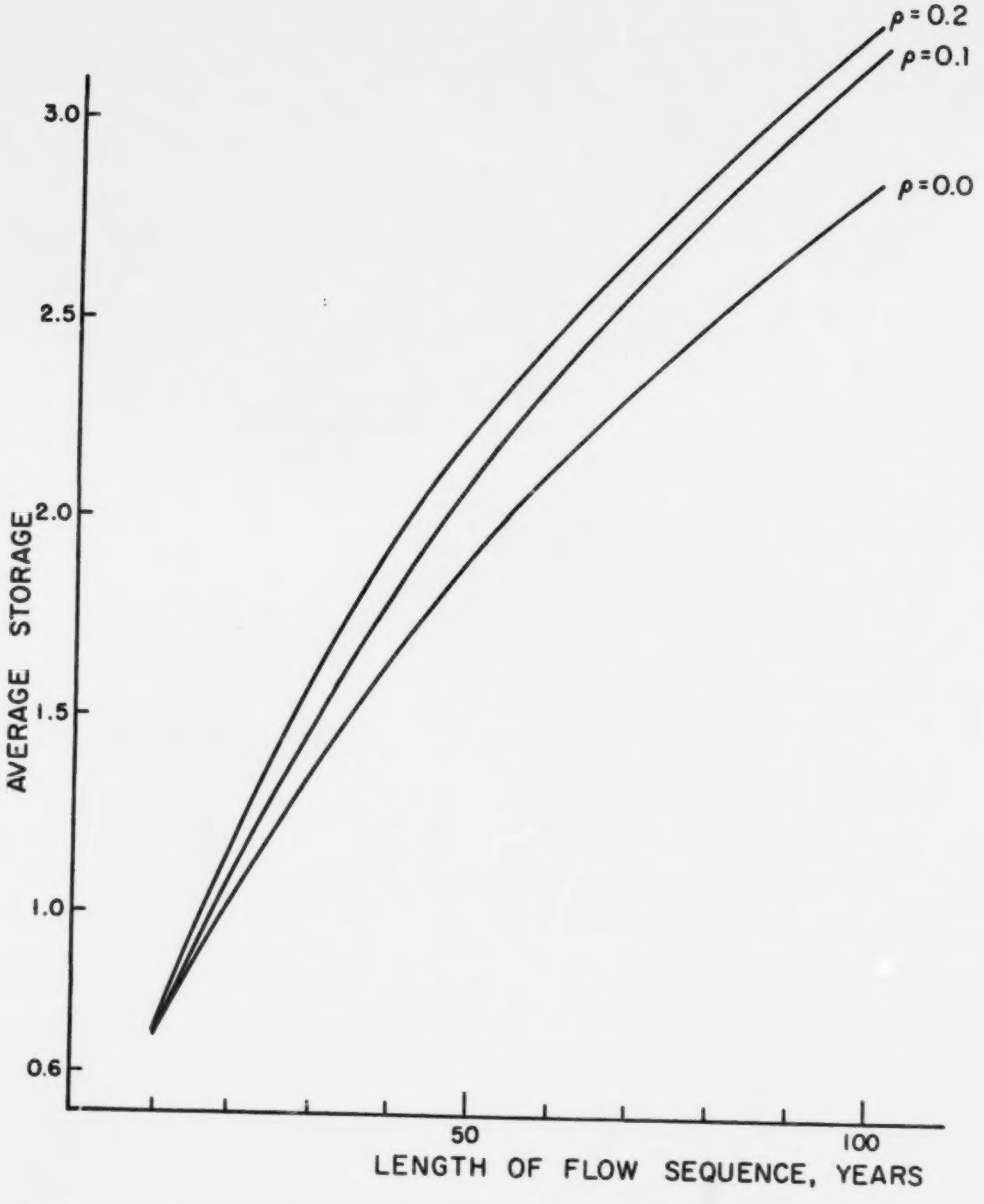
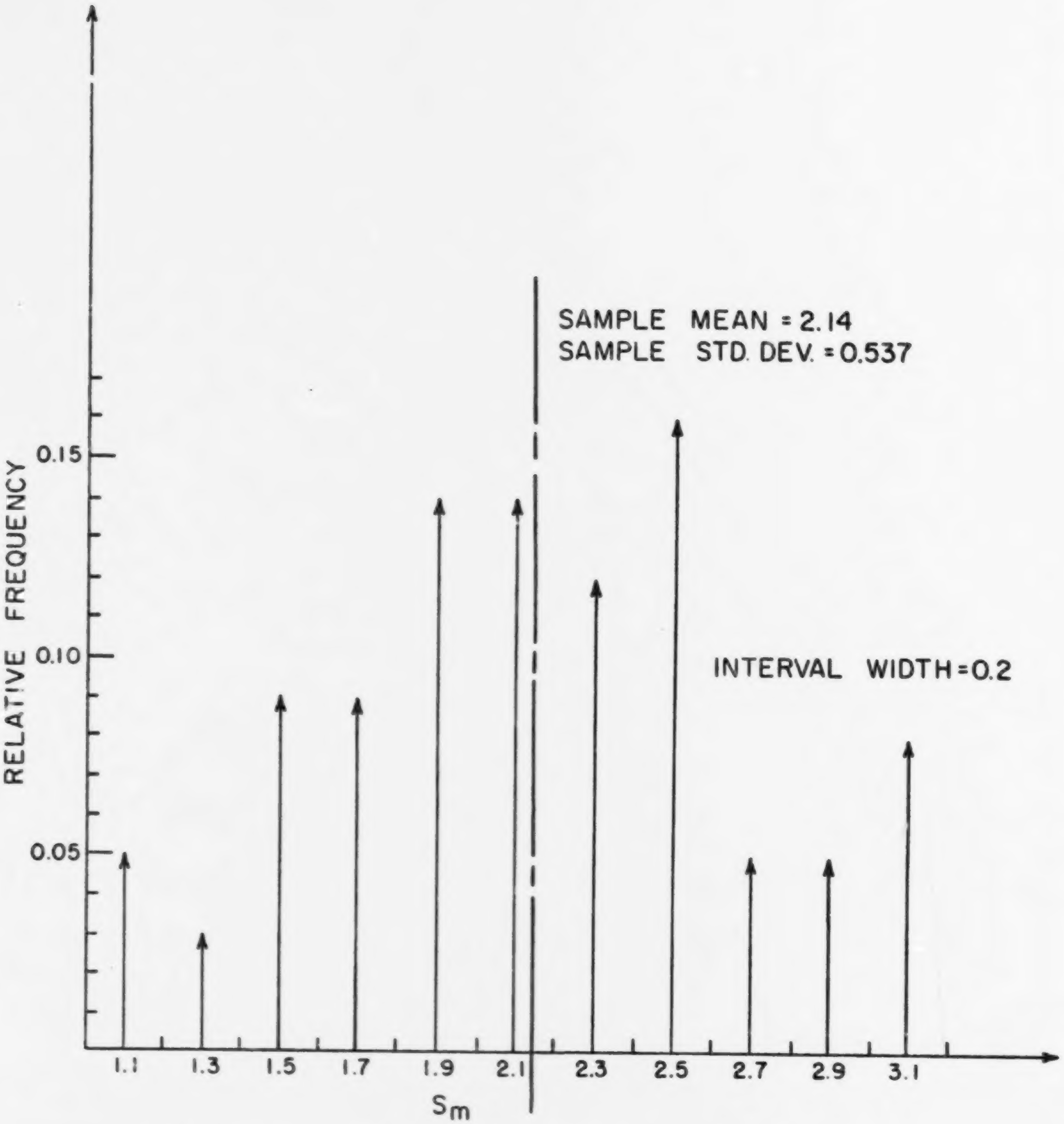


FIG. 4 AVERAGE REQUIRED STORAGE vs. LENGTH OF RECORD, GAMMA DISTRIBUTION ($\gamma = 1.0$), 100% REGULATION

1-26



N = 50 YEARS, 100% DEVELOPMENT, $\rho = 0.1$, $C_v = 0.25$

100 REPLICATE SETS

FIG. 5 DISTRIBUTION OF REQUIRED STORAGE

CHAPTER TWO

Mathematical Model for Computer Simulation of Multipurpose Water Resource Systems

General

Mathematical analysis is the principal and most general tool of the applied scientist for preduction and estimation of the performance of engineering systems. However, in certain complex systems such as multistructure, multipurpose water management systems, the number of relevant variables and their implicit, non-linear interdependencies may be such that a deductive mathematical formulation cannot properly relate the externally specified variables (input) to the internally generated variables (output). Moreover, even if the proper relations can be specified, there may not exist appropriate analytical techniques for surmounting the formidable mathematical complexities which often result. This is often the case in solving non-linear partial differential equations when the boundary conditions are time-dependent or complicated by other demands of realism. When these difficulties arise, as they often do in the analysis of a river-basin system, some form of computer simulation affords a powerful tool for extraction of relevant relations. Simulation analysis is one aspect of the study of mathematical models, and in conjunction with other techniques (steepest ascent, response surface sampling) is the technique whereby we propose to specify an optimal or near-optimal plan for river basin development.

We begin by constructing a model of the system being studied, the model being a representation of the prototype incorporating those features of the prototype deemed to be important for the purpose at hand. We deal with digital models, although analogue and scaled physical models are other familiar representations in engineering practice. The operation of the model in any time interval is characterized by the throughput of input or exogenous objects (water, cars, etc.) in accordance with the parameters of a control mechanism (dams, traffic signals, etc.), yielding a sequence of output or endogenous events and a set of variables which describes the new state of the system. This description of the terminal state becomes the initial state for the following time interval. In certain cases, systems of finite difference or differential equations can be solved to give general analytic solutions for the relationship between exogenous and endogenous events.

If the system is affected or perturbed by random or stochastic components implicit in the sequence of exogenous events, the regulatory mechanism will generally be only partially effective in meeting the objectives of the system. Moreover, the analytic solutions may become so difficult that only application of a simulation technique offers hope of solution. Any single simulation run yields a specific solution, related to the initial conditions of the system, the parameters and nature of the regulation mechanism, and the sequence of exogenous events. To determine the general relation of the exogenous events to the important endogenous events and variables may require a very large number of simulation runs; even then induction from specific results to general solutions will be required.

Monte Carlo techniques can be combined with simulation analysis to remedy the lack of generality implicit in simulation solutions. "Monte Carlo" is the name given to a mathematical technique applied to a problem for which a stochastic process may be adapted which has characteristics satisfying the relations of the original non-probabilistic problem. Such combined studies are used in economic evaluation of simulation results by specifying a distribution of possible endogenous events rather than a single sequence. Decision theory techniques can be applied to this distribution so that the system or component can be ranked in terms of some acceptable metric. Sensitivity analyses can be carried out on the model by performing many runs and changing the value of one or more input parameters each time. The resulting variations in endogenous variables can be observed and related to the corresponding input alterations. The investigator is then better able to decide where to concentrate additional research effort in parameter estimation since wide output fluctuations should be reduced by suitable selection of input parameters.

Until modern digital computers were available, the magnitude and complexity of simulation problems which could be coded were severly limited by the size of available computer memories and the speeds of internal computation, memory access, and input-output. The new IBM 700-series of computers, for example, is sufficiently large and fast to give operational status to solutions of water quality management problems by simulation. The FORTRAN language, an algebraic compiler for IBM computing equipment, provides the link whereby engineers and planners can readily translate their equations and logical operations into the binary language of the computer.

A computer program for simulating the Delaware River Basin is being written as the joint product of our project and the Water Resources Planning Project at the Graduate School of Public Administration. This program includes a subprogram for simulation of the Lehigh River Basin; this subprogram is almost complete and, as the result of exhaustive testing, is free from program errors. Simulation studies of the multi-purpose system involve the following uses: water supply, dilution-flow, recreation, flood control, and water-power. At a later stage when the subroutine for pollution-transport has been completed, investigations will be made of thermal pollution and other water quality problems in the lower reaches of the stream near Allentown and Bethlehem.

The system involves six possible reservoirs, nine possible turbines, two possible aqueducts and one possible diversion channel. Monthly water supply or water quality improvement output is measured in terms of minimum flow maintained in Bethlehem. Monthly energy output is computed as the sum of outputs of the individual power plants. Monthly recreation output is reckoned in the number of visitor-days. Monthly losses in recreation output are measured in terms of reduction of water levels below target levels associated with specific numbers of visitor-days per month. Flood control output is measured in terms of reduction of peak flows at three control points -- Walnutport, Allentown, and Bethlehem, and in the reaches immediately below three reservoirs (Beltzville, Aquashicola, and Trexler).

The entire program is written in the FORTRAN language for the IBM 7090 Computer, and the Lehigh River subprogram utilizes approximately 10,000 binary instructions.

A major effort has been directed at making the operating decisions as flexible as possible; this is accomplished in the monthly operating rule package of six subroutines which comprises the largest segment of the Lehigh River subprogram. Releases are of two kinds: inevitable releases, which are made whenever the sum of initial reservoir content and inflows exceeds available storage capacities, and supplementary releases, which are made whenever additional flow is required to meet any one of several targets (e.g., for energy, water supply, dilution, fish and wildlife). Except by varying the storage capacities, the program cannot alter the inevitable releases. But the supplementary releases can be drawn, in general, from any combination of reservoirs. The computer is coded to recognize a constraint level for each reservoir below which additional supplementary releases cannot be made. The constraint at each site defines a vector in that it can attain a different value for each month of the year. In this way the computer can be instructed to honor such seasonal limits as recreational levels in the summer months and turbine capacities in months of deficient energy production. If the target flows cannot be met when all the reservoirs are drawn down to their constrained levels, the program then accepts a secondary constraint at each reservoir. Additional releases are made down to these new levels (if different from the old), if the targets remain unmet a shortage is recorded and the simulation continues, since no releases can be made below the second constraint. Thus, by varying the constraint levels, priority can be assigned to any of the several uses; consequently, a general flexible operation is attained.

One further use of the variable constraints is being coded at this time. Using available meteorologic data for each month, the program forecasts flows for several months ahead. Three cases are recognized with respect to low flows: safe, unsafe, and critical.

In the first case, no action is taken. But for cases 2 and 3, it may be desirable to hedge against severe drought by accepting small shortages (i.e., losses) during the current month. The variable constraint levels are adjusted to impose such a shortage, the magnitude of the adjustment being related to the severity of the predicted droughts.

There are six subroutines in the monthly operating procedure program: Subroutine BASIC handles the initial water book-keeping, computing "inevitable" flow releases that occur whenever the inflows plus initial contents exceed available storage capacities of the reservoirs, and computing also the releases required to assure minimum flows below the several reservoirs.

Subroutine DIVERT handles diversions of water from the main stem to Mahoning Reservoir.

Subroutine POWER computes the energy generated by the basic and supplementary releases made from all reservoirs. In this program a supplementary release is any release required because the inevitable releases are insufficient to meet the several targets.

Subroutine WATER computes the supplementary releases required to meet target output for water supply at Bethlehem.

Subroutine SUPPOW computes the supplementary releases required to meet the target output for energy.

Subroutine SPACE allocates required releases among reservoirs except those at Bear Creek and Tobyhanna so as to minimize the probability of wasteful spills at the end of the filling season.

The flood generating subroutine utilizes a linear regression of monthly peak on monthly discharge, with an appropriate random component, to compute the synthetic peak flow in any month. By means of a set of estimating equations, the essential features of a flood hydrograph are synthesized, whereupon the peak flows at Bethlehem, Allentown, and Walnutport are evaluated. If the peak exceeds the corresponding channel capacity at these sites or at the upstream damage areas, the flood routing subroutines are called. The hydrographs for the upstream stations provide input for the respective dams, and the maximum releases which cause no damage at the upstream channels or the downstream damage areas are made. If the upstream reservoirs fill, the spills are routed downstream to the flood damage areas and there are compared to the natural flows estimated from the synthetic hydrographs. Reductions in flood damage are computed as a function of the reductions in peak flows.

The studies described here will be used in an effort to delineate the optimal plan for development of the Delaware River basin. At the start an objective function must be specified; benefits are generated by the system outputs only when these outputs are directed toward satisfying the objective function. Then a set of input variables is selected, either by a random or systematic sampling scheme, and the system thus specified is simulated on the computer. The resulting benefit defines a point on the multi-dimensional response surface.

Many such points are located, and the system which produces the most favorable response is selected as the optimal solution. If m random combinations of inputs are selected, it can be shown that the probability that at least one of the resulting benefits will lie within the upper p% of the population of benefits is $1-(1-p/100)^m$. No assumption as to the nature of the distribution of benefits need be made. Thus, if 30 trials are made, the probability that at least one observation (benefit) falls in the upper 10% of all benefits is $1-(1-10/100)^{30}$ or 0.957. The usefulness of this result depends on the peakedness of the response surface, since the upper 10% of all benefits will encompass a small or great range of values depending on whether the topography of the response surface is gentle or jagged. Futher refinements in optimization, such as steepest ascent techniques, are being studied and will be applied to the model as they are perfected.

Program for the Monthly Operating Procedure, Lehigh River

Introduction

To help the reader follow the detailed presentation of the computer logic, we first discuss an idealization of the Lehigh river system; this is followed with a general discussion of the over-all simulation program, which shows how the monthly operating program package is related to the total program.

The detailed presentation of the program starts with a brief statement of the general operating policy of the monthly program, and continues with a step-by-step statement of the program in detail, supported by flow charts which show the logical sequence of the simulation in FORTRAN and other standard symbolic terms. A glossary of FORTRAN symbols is appended to the chapter to enable the reader to interpret the flow charts. The chapter concludes with a discussion of problems encountered in preparing this part of the simulation program and suggestions for practical application of the technique; and with observations and questions on the relation of our experience to constructing and programming optimal operating procedures.

The Lehigh Basin System

Figure 1 presents the stream and structure configuration of the Lehigh System. There are a total of six reservoir sites to be identified throughout the simulation program by the numbers 1-6. Reservoirs 1 and 2 (Tobyhanna and Bear Creek) are alternatives; thus, any system design can include one of these reservoirs but not both. In effect, we deal with alternative system designs which can include up to five reservoirs.

Each reservoir can be used to provide (1) regulated flows (for water supply or water quality improvement) at Bethlehem, (2) recreation facilities at the reservoir site, (3) flood control storage, and (4) hydroelectric energy.

As shown in Figure 1, a diversion channel can be provided from the Lehigh River at a point below Reservoirs 1 and 2 to Reservoir 3 (Mahoning) so that flows in the Lehigh can be diverted and stored in Reservoir 3. This reservoir thus functions as an off-channel storage unit.

Reservoirs 3, 4, 5 and 6, can each have but a single hydro power plant location. But, as shown in Figure 1, Reservoirs 1 and 2 can have more than one (although only one for each design). Actually, there are five variations, which are identified as cases in the detailed program; these exist in conjunction with Reservoirs 3-6.

The associated turbine assignments are:

Case	Reservoir No.	Turbine No.*	Location of Power Plant
1	1	1	At the dam site discharging into the Lehigh directly below the dam site. Called "no-pipe"
2	1	7	At end of a high-head penstock, discharging into Reservoir 4 (Beltzville). Called "Beltzville pipe"
3	2	2	At the dam site discharging into the Lehigh directly below the dam site. Called "no-pipe"
4	2	9	At the end of a high head penstock, discharging into the Lehigh, several miles below the dam site. Called "Lehigh Pipe"
5	2	8	At the end of a high-head penstock discharging into Reservoir 4 (Beltzville) Called "Beltzville pipe"

Selection of a specific system design implies selection of a single case.

In all, there are 36 major decision variables. These are:

Variables	Units	No. of Variables
Reservoir storage, allocation to dead storage	10^2 ac ft	6
Reservoir storage, allocation to flood storage,	" " "	6
Reservoir storage, total capacity	" " "	6
Power plant capacities	Megawatts	9
Capacity diversion works, Lehigh to Res. 3,	10^2 ac ft/ month	1

---

* Turbine Nos. 3-6 are assigned to Reservoirs 3-6

Variables	Units	No. of Variables
Target output for energy	megawatt hours/month	1
Target reservoir level for recreation, expressed as a range	feet above stream bed	6

In a broad sense, all of these can be considered as parameters of the operating procedure and can be changed at preset times during the simulation. Decisions on total capacities of reservoirs, capacities of power plants, and diversion works require fixed commitments for capital investment and would not be changed often during a run. Allocations to dead storage are generally fixed, because power plant design requires a fixed level of dead storage. While allocations to flood storage can be allowed to vary cyclically from month to month, the minimum allocation over the simulation period must be fixed for each design, as this sets the approximate level of the reservoir pool below which recreational structures and facilities would not be provided. Target outputs for water supply and energy and target reservoir levels for recreation can be allowed to vary from month to month and over the years, but there is some relationship between the maximum monthly target output for energy and size of power plants, as well as between maximum monthly target reservoir levels and location and nature of recreation structures bordering the reservoir.

Associated with these major variables are lesser ones, such as penstock capacities and heads, power plant efficiencies and maximum heads, which are related to the major variables. In addition, the monthly operating procedure recognizes 6 minimum flow constraints -- one for each reservoir -- specifying the minimum flows to be provided in the streams below the reservoirs for amenities and for maintenance of fish life. These constraints are likely to be of minor significance because so little water is required to meet them. Of more importance are the parameters that control the apportionment of releases from Reservoirs 3-6 by means of the space rule. These are now expressed as the ratio of expected inflow to each reservoir for the remainder of the drawdown-refill period to the total expected inflow to all four reservoirs over this period; the four ratios must have a sum of unity. By redefining these factors, different apportionments of releases can be effected.

The Over-all Simulation Program for the Lehigh System

Given this system of six possible reservoirs and nine possible turbines with the configuration of Figure 1, we are constructing a simulation program to study the behavior of any given design over an extended period of time. The heart of the program is the monthly operating procedure. At present we visualize the over-all simulation program to be related to the monthly operating procedure as shown in Figure 2.

With specified values for the design variables and initial conditions, the Executive Program calls the Monthly Streamflow Generating Program to produce the inflow hydrology for the first month. (Alternatively, historic inflow hydrology can be provided as input.) Control then passes to the Monthly Operating Program which operates the system for the first month, as described in detail in the sections that follow. The flood flow generating program is then called to determine if damaging flood flows are experienced. If so, a synthetic flood hydrograph is constructed at each site and the flood routing program is called. Historic flood hydrographs can be introduced, if desired. If there are no damaging flood flows, the low-flow generating program is called to determine if critical low-flow conditions are encountered. If so, the low-flow hedging rule is introduced in an effort to reduce future catastrophic droughts by taking small current losses. This completes the operation for the first month; the program then processes the second and succeeding months of the simulation period until the run is completed. At the end of the simulation period, control passes to the benefit sub-program which evaluates the performance in monetary terms and discounts the time-streams of costs, gross benefits, and net benefits to present value. The output reporting program is called to record the desired output for off-line printing.

The Monthly Operating Program in Detail -

Basic Policy

The basic policy is, first, to provide for certain specified minimum flows below the reservoirs and, second, to meet the target outputs for water supply to Bethlehem and electric energy for the system, subject to specified constraints on minimum and maximum reservoir contents.

Inflows that cannot be stored within the constrained reservoir range are used to generate electric energy, subject to the available capacity of penstocks and turbines, and contribute toward meeting the water supply target at Bethlehem. Stream flows in the Lehigh at the point of diversion to Mahoning Reservoir (No. 3) that are not needed to meet the target water supply flow are diverted to Reservoir 3, subject to the limitations of the diversion channel or space available in Reservoir 3.

Reservoir releases required to meet the target water supply are made first from Reservoir 1 or 2,* subject to a variable constraint on minimum reservoir contents. If necessary, additional releases are made from Reservoir 3, 4, 5 and 6, according to a rule which apportions the releases so that the total remaining storage space is divided among the reservoirs in accordance with specified proportions.

These proportions are chosen to equalize the probability of spill during the remainder of the drawdown-refill cycle among the four reservoirs, under estimated patterns of future inflows.

---

* Because these reservoirs are alternatives, only one will appear in any trial system design.

If further releases are necessary to meet the target water supply flow, Reservoir 1 or 2 is again utilized with a revised, lower constraint on reservoir contents. If further releases are required, they are made from Reservoirs 3, 4, 5 and 6, according to the rule described above, but with a revised (lower) constraint level on reservoir contents.

All releases for water supply are routed through the hydroelectric plants up to the limits of penstock and turbine capacities, but if these releases fail to generate sufficient energy to meet the target output, additional releases are made in the same sequence as for water supply.

If target outputs for water supply or energy are not met after releases are made from all reservoirs to the full extent allowed by constraints on storage level and on penstock or turbine capacities, no further releases are made; instead, the deficits in water supply or energy for the month are recorded. In any event, the monthly operating policy is terminated at this point, and control is returned to the Executive Routine for investigation of floods and low flows.

The Subroutines in Detail

As shown in Figure 3, the Monthly Operating Program consists of six subroutines: BASIC, DIVERT, POWER, WATER, SUPPOW, and SPACE. The following sequences describe the several subroutines in detail; Figures 4-9 provide further clarification. The calling sequence appears in Figure 3.

Subroutine BASIC (Figure 4)

1. For each reservoir, define the basic release (+)[1] as the sum of inflow and initial contents minus the allowable storage capacity (i.e., the total storage capacity less the flood storage allowance).

---

[1] In this discussion, the symbol (+) will be taken to mean the stated quantity if this quantity be positive; otherwise the quantity is modified so that its value is zero.

2. For Case 2, define the excess flow (+) as the difference between the basic release and the minimum channel flow required below Reservoir 1, this difference limited by the Beltsville penstock (or more simply, the penstock) capacity. The computation is similar for Cases 4 and 5, except that the excess flow (+) is defined with respect to Reservoir 2.

3. For Cases 2 and 5 redefine the basic release at Reservoir 4 by adding to the previous value the newly defined excess flow. This has the effect of routing all of the water which is diverted from Reservoir 1 (or 2) through Reservoir 4, storing none; if necessary, this is subsequently modified in Subroutine POWER.

4. For each reservoir, define the shortage as the difference between the minimum flow required at the reservoir and the basic release. If this shortage is not positive, the preliminary computation is completed for the reservoir; if the shortage is positive, the first supplementary release, limited by the water available in the reservoir, is made so as to meet the minimum flow requirement.

5. The program returns to the EXECUTIVE Program, thence to Subroutine POWER.

Subroutine POWER (Figure 5)

1. For Reservoirs 1-6, update the contents by computing the sum of the inflow and the initial contents, and subtracting from this the sum of the basic release and the supplementary releases made to date.

2. For Reservoir 3, increment the new (i.e., updated) contents by the diversion from the main stem. The first time that Subroutine POWER is called, no diversion has been defined so that this computation gives no change in contents. Define, as the additional supplementary release available for use at Reservoir 3, the sum of the second and third supplementary releases made to date from Reservoir 1 (or 2).

3. For Reservoir 4, Cases 2 and 5, the computation proceeds as follows. (For Case 2 read Reservoir 1 and for Case 5, read Reservoir 2.) Define the penstock flow as the sum of the excess flow and the second and third supplementary releases from Reservoir 1 (or 2), this quantity limited by the penstock capacity. If no supplementary releases are required at this stage of the computation, zero is substituted for the relevant addends. In the event of a flow deficiency at Bethlehem, the first supplementary release at Reservoir 4 is incremented by the total flow through the penstock: i.e., all the water arriving at the reservoir is routed directly through the turbines and thence downstream. Using these data the basic release (+) and the updated contents at Reservoir 4 are recomputed. The new contents are compared to the normal pool contents to ascertain that penstock inflow has not raised the storage level so as to impinge on the space required for flood control; if so, additional supplementary releases are made. In the event that the penstock capacity limits the flow into Reservoir 4, the unused water is routed directly downstream through the channel below Reservoir 1 (or 2).

4. Having updated all the reservoir storage values, the program proceeds to evaluate the energy produced by the several releases at each reservoir. The average storage, being one-half the sum of the initial and current storages, is computed. The total time, expressed as a fraction of a month during which the contents are at or above the dead storage level is computed, utilizing the assumption that the reservoir contents change linearly during the course of the month.

5. The average head at each reservoir is evaluated by means of a fourth-order polynomial function of contents, not all of whose terms need be utilized. The average heads corresponding to turbines 7, 8, and 9 are computed by adding the penstock drop to the appropriate reservoir head.[1/]

[1/] If pumping is assumed from Reservoir 1 (or 2) heads for turbines 7-9 are fixed.

6. The water capacity at each of the 9 turbines is evaluated as a function of average head and turbine characteristics (i.e., power capacity and maximum head).

7. The effective power flow at turbines 4-6 is computed as the sum of the basic release and the three supplementary releases. The effective power flow at turbine 3 is computed in a similar fashion, except that the additional supplementary release (step 2) is included in the sum.

8. For Case 1, the effective power flow at turbine 1 is computed as in step 7 above; the computation is similar for the effective power flow at turbine 2, Case 3.

9. For Case 2, the effective power flow at turbine 7 is given by the sum of the excess flow and the second and third supplementary releases from Reservoir 1; that is, by all the water released from Reservoir 1 except that which is required for a minimum flow downstream of the site.

10. For Cases 4 or 5, the effective power flow at turbine 9 (or 8) is given by the sum of the excess flow through the penstock and the second and third supplementary releases from Reservoir 2.

11. The effective power flows are limited for all 9 turbines by the respective turbine capacities.

12. The energy produced at each of the 9 turbines is computed and accumulated, whereupon the total energy production is compared to the energy target.

13. If the target is met or exceeded, the energy deficit is set equal to zero and the dump energy is computed. If the target is not met, the dump energy is set equal to zero and the energy deficit is computed.

14. The routine returns to the calling location, i.e., to the instruction sequence in whichever subroutine called Subroutine POWER.

Subroutine DIVERT (Figure 6)

1. The unregulated flow for the reach of river between the reservoirs and Bethlehem (the intermediate flow) is computed.

2. The regulated flow at Bethlehem, being the sum of the intermediate flow and all the basic and supplementary releases, is computed.

3. The Bethlehem flow is compared to the water supply target, whereupon 3 branches are available.

4. If the Bethlehem flow equals the target the amount of water available for diversion and the water deficit are set equal to zero. Go to Step 13.

5. If the Bethlehem flow is less than the target, the deficit is computed and control passes to Subroutine WATER for the allocation of the second supplementary releases among the reservoirs. Go to Step 13.

6. If the Bethlehem flow exceeds the target, the deficit is set equal to zero and the amount of water available for diversion is computed according to the following scheme.

7. For Case 1, the amount available is the natural inflow between Reservoirs 1 and 3 plus the basic release and first supplementary release (for minimum channel flows) from Reservoir 1.

8. For Case 2, the amount available is computed as in Step 7 except that the effective power flow through turbine 7, or that quantity which is released to Beltzville, is subtracted.

9. For Cases 3 and 4, proceed as in Step 7 except that Reservoir 2 is applicable where Reservoir 1 appears.

10. For Case 5, proceed as in step 7 and 8 except that Reservoir 2 and turbine 8 are applicable where Reservoir 1 and turbine 7 appear.

11. The amount actually diverted is limited by the space available in Reservoir 3, the capacity of the diversion works, the excess of flow over

target, or the amount available for diversion, whichever is smallest.

12. The contents in Reservoir 3 are incremented and the flow at Bethlehem is decremented by the amount diverted.

13. If the energy deficit is ≤ 0, control passes to the EXECUTIVE Program; if the energy deficit exceeds zero, Subroutine SUPPOW is called to allocate the set of third supplementary releases among the reservoirs, whereupon control returns to the EXECUTIVE Program. The monthly procedure is complete when the program returns to the EXECUTIVE Program from this subroutine.

Subroutine WATER (Figure 7)

1. For each reservoir, define the surcharge (+), or the quantity of water available above the first (i.e., upper) variable constraint.

2. For Cases 1 and 2, attempt to meet the water deficit by making the second supplementary release from Reservoir 1. This release is limited by the surcharge at Reservoir 1.

3. For Cases 3, 4 and 5, proceed as in step 2 except that Reservoir 2 should be inserted where Reservoir 1 appears.

4. If this supplementary release is sufficient to meet the water deficit at Bethlehem, call Subroutine POWER to evaluate the effects of the release and then return control to Subroutine DIVERT.

5. If additional releases are required, define the new requirement so as to provide an input for Subroutine SPACE, whereupon control is transferred to that subroutine and allocations are made from Reservoirs 3-6. The new flow at Bethlehem is computed and, if the water supply target is met, control is transferred to step 4.

6. If the water supply target is not met, steps 1-5 are repeated, using the second (i.e.,lower) constraint to redefine the surcharge (+). If the water supply target cannot be met, a shortage is noted, Subroutine POWER is called, and control is returned to Subroutine DIVERT.

Subroutine SUPPOW (Figure 8)

1. Call Subroutine POWER to evaluate the power generated by all the releases in the system. If the energy deficit is ≤ 0, return control to Subroutine DIVERT.

2. For each reservoir, define the surcharge (+), or quantity of water available above the first variable constraint.

3. For each turbine, define the unused turbine capacity as the difference between the water capacity and the effective power flow.

4. For each reservoir, compare the surcharge to the unused turbine capacity. If the surcharge is the greater of the two, increase the constraint level by the difference, thereby decreasing the surcharge to the level beyond which additional releases would produce no incremental power.

5. For Cases 2, 4 and 5, the surcharges in Reservoir 1 or 2 are limited by the unused capacities of turbines 7, 9 and 8, respectively; any additional release is passed downstream in the channel, bypassing the turbines.

6. Based on the energy deficit, a trial supplementary release is made from Reservoir 1 (or 2). This release is routed through turbine 3, so the conditions in both Reservoir 3 and Reservoir 1 (or 2) must be examined to arrive at a reasonable approximation to the required flow. This release is limited by the surcharge in Reservoir 1 (or 2).

7. Call Subroutine POWER to evaluate the effects of the release made in step 6.

8. Limit the surcharge in Reservoir 3 by the newly computed unused turbine capacity, taking account of the release defined in step 6.

9. Make the energy test: i.e., determine if the generated energy meets the target within a specified tolerance. If so, return control to Subroutine DIVERT, thence to the EXECUTIVE Program.

10. If the energy deficit is negative, implying too large a release, reduce the release accordingly and go to step 7.

11. If the energy deficit is positive, release the entire surcharge in Reservoir 1 (or 2). Go to step 7 and evaluate the effect of this release.

12. If the energy deficit is still positive, additional releases are required from Reservoirs 3-6. Compute the average head and average time (at or above dead storage) for these reservoirs, and evaluate a trial requirement whereby the energy deficit will be satisfied.

13. Call Subroutine SPACE and allocate releases among Reservoirs 3-6.

14. Compute updated surcharge (+) values for Reservoirs 3-6.

15. Call Subroutine POWER and compute the energy generated by the system.

16. If the energy generated meets the energy target within the specified tolerance, return control to Subroutine DIVERT.

17. If the energy deficit is negative, restore the original state of the system, compute a new (reduced) requirement, and go to step 12.

18. If the energy deficit is positive, count the number of reservoirs for which the surcharge exceeds zero (i.e., from which additional releases may be made).

19. If the count exceeds one, compute a new average head and go to step 12 to compute a revised requirement.

20. If the count is one, try to make the required release from the corresponding reservoir, noting that the magnitude of the release is limited by the surcharge.

21. Call Subroutine POWER and compute a total energy generated.

22. For each reservoir, reduce the surcharge to take account of the latest release.

23. If the energy generated meets the energy target within the tolerance, return control to Subroutine DIVERT.

24. If the energy deficit is negative, reduce the requirement and restore the original state of the system; then go to step 21.

25. If the energy deficit is positive, all the reservoirs are now drawn to their constraint levels; it is necessary to replace the first constraint with the second and re-execute steps 2-24.

26. If an energy shortage still exists, note the shortage and return to Subroutine DIVERT.

Subroutine SPACE (Figure 9)

1. Note the basic release allocation equation:

$$\text{Allocation}_i = \text{Contents}_i - \text{Capacity}_i + \frac{+ [\Sigma(\text{Capacity}_i - \text{Contents}_i) + \text{Release Requirement}] \cdot \text{Future Inflow}_i}{\Sigma \text{ Future Inflows}_i}$$

where i is the reservoir number and the summations are performed over $i = 3, 4, \ldots, 6$.

2. Compute Σ_i Future Inflows$_i$ and Σ_i Capacity$_i$, and note the magnitude of the externally specified requirement.

3. Compute relevant surcharge (+) values, depending upon the nature of the requirement (i.e., whether there exists an energy or water deficit). Note that the effects of turbine and variable constraints are considered in the calling routine, not in this subroutine. That is, Subroutine SPACE may be used to make allocations for any purpose.

4. Update all reservoir contents calculations as in steps 1-3 of Subroutine POWER.

5. Count the number of reservoirs in which the surcharge exceeds zero.

6. If the count is zero, return to the calling location since no gain can be effected by continuing in this subroutine.

7. If the count is one, make the required release from the corresponding

reservoir, up to the surcharge limit, and decrement the contents accordingly. Then return to the calling location.

8. If the count exceeds one, test the feasibility of further computation by summing the 4 surcharges and comparing the sum to the requirement.

9. If the requirement ≥ the sum of surcharges, release all the surcharges and decrement the storage values. Then return to the calling location.

10. If the sum of surcharges exceeds the requirement, control will remain in this subroutine until the requirement is satisfied. Recompute Σ_i Contents$_i$ for use in the basic allocation equation.

11. For each reservoir in turn, starting with Reservoir 3, test to determine whether the reservoir has been removed from the set of admissible sources of water. The first time this step is executed, all reservoirs are admissible.

12. If the reservoir has been removed, increment the reservoir number and return to step 11 until all reservoirs are completed; then go to step 19.

13. If the reservoir is admissible, test to determine whether it is the only admissible reservoir.

14a. If only one reservoir is admissible, test to determine if the surcharge therein is adequate to meet the requirement. If so, go to step 7 to make the necessary releases. If not, release up to the limit and return to execute step 3 only. Control is then transferred to step 14b.

14b. If all reservoirs are drawn down to their constraint levels, return to the calling location.

14c. If one reservoir remains unutilized, go to step 7 to withdraw the required surcharge.

14d. If more than one reservoir remains unutilized, go to step 11 and restart the allocation process.

15. If no reservoirs are admissible, recompute the surcharge for all 4 reservoirs and begin the computation anew by making releases from those reservoirs which have positive surcharge values. This step is rarely reached: indeed, certain extreme combinations of input data are required to induce the enormous inbalance in allocations which are implicit in executing this step. Finally, return to the calling location.

16. If more than one reservoir is admissible, make allocations in accordance with step 1.

17. If an allocated release exceeds the surcharge, release up to the surcharge. Then reduce the requirement and update the contents; this reservoir, fully utilized, must now be withdrawn from the set of admissible reservoirs. Σ_i Future Inflows$_i$, Σ_i Capacity$_i$ and Σ_i Contents$_i$ are recomputed, and the allocation begins anew at step 11.

19. After all allocations are made, their sum is compared to the original requirement to ascertain that no continuity error has been made. If an error is detected, the routine stops and requests a memory dump. If no error is detected, the routine returns to the calling location.

Lessons for Practical Application of the Technique

1. Effective application of simulation programming requires ready access to a high speed digital computer with a large memory capacity - for example, the IBM 700-series. River system simulations are complex and involve manipulation of enormous amounts of data, so that the storage capacity of even the largest machines is soon taxed. For example, our monthly operating package alone requires the use of 15,000 of the more than 32,000 memory cells in the IBM 7090.

2. Effective operation requires that an algebraic compiler be available for the computer so that intuitively palatable mnemonics can be used by

hydrologists and engineers instead of machine language. FORTRAN and ALGOL are two such compilers.

3. As for the details of programming, simple checks - such as the maintenance of continuity - are not nearly so straightforward as one might expect. The detailed flow relationships are not algebraically trivial, and subtle inconsistencies in the routing of flows are often built into the program. Frequent intermediate checks and core dumps are the only means to uncover these errors. Complications in programming and testing arise more rapidly for reservoirs in series than for reservoirs in parallel; this follows from the interdependence of flow patterns for reservoirs in series.

4. Testing of programs should be made with polar cases in order that errors and inconsistencies can be spotted more readily.

5. When several different cases - that is, configurations of flow - must be dealt with, it is probably better to code them as independent operating programs than to combine them in a single master program, as we have done. Each case can then be dealt with independently, and the chore of coding and debugging can be eased.

6. As might be expected, the simple bookkeeping transactions, such as adding inflows and subtracting outflows, are the most straightforward and the easiest computations to program. FORTRAN is particularly powerful in compiling these algebraic and transcendental functions, such as those which occur in computing energy generation.

7. In contrast, the most difficult problems arise in treating logical statements. We found the space rule to be particularly difficult to program because of the many alternative decisions imposed by the constraints.

8. The program for the space rule as now written is perfectly general. It can handle any number of reservoirs, and constraints can be set at any

level of reservoir content. The parameters that govern the apportionment of reservoir space (and, hence, releases) can also be changed at will.

9. At this point in the research, we are not prepared to define the limits on the size of simulation problems that can be handled effectively. The simulation of a river system as large and complex as the Delaware is a major task - one that strains the resources of our group. Yet, it would appear to be perfectly feasible for river systems of this and even larger size to be simulated by a planning staff composed of appropriate numbers of skilled engineers and programmers.

10. Programming of river system simulations will typically be so time-consuming that more than one person will be needed to analyze, code, debug and test the program. Communication among the participants will be much easier if mnemonics are used than if resort must be had to machine language. For effective programming the problem analyst and programmer must have a background in river basin planning. Simulation of river systems cannot be assigned entirely to professional programmers without guidance from river system planners. It appears that a relatively small staff of river system planners and programmers would be adequate to the programming task.

Relation to Optimal Operating Procedures

The operating procedure that is built into the monthly program makes two kinds of choices: (1) choice of releases for different purposes and (2) choice of releases from different reservoirs. As for the first, flood control and recreation are introduced into the simulation by setting monthly constraints on reservoir levels. In a sense, these purposes can be said to have first priority, since other releases must conform to these several constraints. Within the month, releases for water supply are given priority over releases for energy. We can determine explicitly the quantity of the release required to meet the water supply target, while the non-linear relations between flow and energy impose an iterative determination of the required release for energy. It is thus simpler to release first for water supply and then determine if further releases are needed for energy than to attempt the converse.

The choice of releases among reservoirs is done in stages as follows: (1) Reservoir 1 (or 2) is first called upon for releases followed, if necessary, by demands on Reservoirs 3-6. (2) Releases required from Reservoirs 3-6 are apportioned by the space rule. Choices (1) and (2) are made subject to a single set of constraints on reservoir level (and volume), and if necessary these choices are repeated for a second set of constraints (in which at least one constraint would be less binding than in the first set).

Reservoir 1 (or 2) is drawn down separately from the others because the outflows from this reservoir can be stored, at least in part, in Reservoirs 3 and 4. Due to this complementarity in storage, inclusion of Reservoir 1 (or 2) in the space rule would add logical and numerical complications. We admit that the decision to call first on Reservoir 1 (or 2) for releases is arbitrary. But, we can easily reverse the order by setting the variable constraint level at its highest possible value for Reservoir 1 (or 2), in effect, making releases first from Reservoirs 3-6, because no releases will be made from the upstream reservoir in the first pass.

It is clear from the above that our procedure has many arbitrary aspects and that we cannot claim optimality. But the procedure is also very flexible, in the sense that an entire family of procedures can be created by manipulating the two sets of reservoir level constraints. By this means, first priority of releases can be assigned either to Reservoir 1 (or 2) or to Reservoirs 3-6. In addition, releases from one or more of Reservoirs 3-6 can be set to zero by an appropriate setting of the constraints. The apportionment of releases among Reservoirs 3-6 also can be varied by changing the manner of computing the space allocations to reservoirs.

The operating procedure will ultimately incorporate a hedging rule. That is, target outputs for the month are met with regard for possible deficits in succeeding months. The procedure does not look ahead to determine if the reservoir system can handle expected inflows without incurring wasteful spills.

No progress can be reported in the development of an optimal operating policy for a river system of this magnitude. And if such a procedure were available, it could not be made a part of the Monthly Operating Package described in this document.

List of FORTRAN Symbols that appear in the Flow Charts, Fig. 4-9*

ACTUAL - final allocation of releases, Res. 3-6, made by the Space rule, 10^2 ac ft.

(6) C SPACE

ADDSR - additional supplementary release from Res 1 or 2 routed through Res 3 turbine - constrained by channel or turbine capacity, 10^2 ac ft.

C POWER

AVAIL - space available for storing water in reservoirs 1-6, 10^2 ac ft.

(6) C DIVERT

AVHD - average head of Res 3-6, used for preliminary computation of the required release, ft.

C SUPPOW

AVTIME - average fraction of a month during which reservoirs 3-6 are at or above the dead storage levels, used to estimate the release required to meet the target output for energy.

C SUPPOW

B - coefficients of the head-capacity or other equations; B(i,j,k) interpreted as the i^{th} coefficient at the j^{th} location, k^{th} use.

(5, 10, 50) C

---

[*] Following the definition of each symbol, there will appear certain technical FORTRAN information coded as follows:

On the left-hand side of the page, the number in parentheses, e.g., (6), denotes the dimension of the matrix, if the variable is subscripted. Absence of the number means that the variable is unsubscripted.

In the center of the page, presence of the letter C denotes that the variable appears in COMMON storage; its absence means that the variable is not in COMMON storage.

To the right, the subroutine symbol, e.g. SPACE, denotes the subroutine in which the variable is first defined. Absence of a subroutine symbol means that the variable is either read as an input or is assumed as an input to the Monthly Operating Package.

BASIC - subroutine of the Monthly Operating Package.

BR - basic release, or inevitable flow necessary to keep contents at flood storage level, for reservoirs 1-6, 10^2 ac ft.

(6) C BASIC

C - coefficient for computing the water available for diversion from the Tannery gauge

C

CAP - reservoir capacity, reservoirs 1-6, in 10^2 ac ft.

(6) C

CAPDIV - capacity of diversion chennel to Res. 3, 10^2 ac ft.

C

DEAD - dead storage, associated with turbines 1-9, 10^2 ac ft.

(9) C

DEFCT - deficit in energy - or target output for energy minus total energy generated, Mw hr.

C POWER

DENOM - sum of future flows (FUFLO) for reservoirs for which space rule allocations are to be made, 10^2 ac ft.

C SPACE

DIVERT - subroutine of the Monthly Operating Package.

DSR - change of supplementary releases 1-3 at reservoirs 1-6, computed by the Space equation. This is constrained to be ≥ 0 and $\leq$ total contents above the constraint, 10^2 ac ft.

(3, 6) C SPACE

DSRR - notation for current incremental supplementary release, 10^2 ac ft.

C SPACE

DUMP - energy produced in excess of target output for energy, mw-hr.

C POWER

EFFPOW - effective power flow, turbines 1-9, 10^2 ac ft.

(9) C POWER

EFFY - efficiency of the power plants.

C

ENERGY - energy generated at each of the 9 turbines, in mwh.

(9) C POWER

ENGTOT - total energy developed in a month, mwh.

C POWER

ERRTOL - coefficient of energy target, for evaluating whether the generated energy is within the accepted tolerance.

C

ERTOL - constant term used in evaluating whether total generated energy approximates target output for energy within tolerance limits.

C

ETARG - target output for energy, mwh.

C

EXCESS - non-negative difference between the inevitable flow and the minimum flow requirement, 10^2 ac ft.

(6) C BASIC

EXTRA - to provide emergency floating point communications between all subroutines.

(1000) C

FACTOR - factor converting c x 10^2 ac ft to Mw-hr in equation E = FACTOR Qhe It has value of 0.1024

C

FLALL - flood allowance at Bethlehem

FMT1(12) - arbitrary format statement for coefficients, read in

C

FMT2(12) - arbitrary format statement

C

FMT3(12) - arbitrary format statement

C

FMT4(12) - arbitrary format statement

C

FNRES - total number of reservoirs to which the space rule is applied

C

FSA - space allocated to flood control storage, reservoirs 1-6, 10^2 ac ft.

(6) C

FSHORT - deficit of actual releases by space rule when compared with required releases, 10^2 ac ft.

C SPACE

FUFLO - future flows (over remainder of the drawdown-refill period), Reservoirs 3-6, corresponding to each month, for use in the space rule, 10^2 ac ft.

(4, 12) C

HARB(3) - heads (incremental) associated with penstocks, ft.

C

HAVG - average head at the turbines 1-9, computed from the average contents, ft.

(9) C POWER

HMAX - maximum head, power plants 1-9, ft.

(9) C

ICASE - case number, 1-5

C

ICOUNT - number of reservoirs at which the current storage is equal to or exceeds the current constrained level.

C SUPPOW

IEXTRA - symbol to provide emergency fixed point communications between all subroutines

(500) C

II - loop control index in SUPPOW.

C

IMONTH - cyclical index, 1 to 12, for months.

C

IN - sum of the reservoirs which have not encountered constrained releases in the Space allocation.

C SPACE

INUSE - number of reservoirs which have not encountered constrained releases in the Space allocations.

C SPACE

IRUN - number of the simulation run

ISPACE - index to inform the SPACE routine of the origin of the current shortage

if ISPACE = 2, water supply is deficient;

if ISPACE = 3, energy is deficient.

POWER

ISTART - month number, 1-12, in which the simulation begins.

C

J - index for reservoirs 1-6

C

JCOUNT - sum of the reservoir numbers for which current contents exceed the current constraint storage.

C SUPPOW

K - index for distinguishing between reservoirs 1 and 2, as a function of the current case under investigation.

C POWER

KCOUNT - number of reservoirs at which current contents exceed the current constrained storage.

C SUPPOW

KK - index for reservoir and turbine listing

DIVERT

KLIM - limit on coefficient input

KM - variable call parameter for "ASSIGNED GO TO" loop in SUPPOW.

C

L - loop counter for SUPPOW.

SUPPOW

LETSGO - variable address for assigned transfer statements.

SPACE

LJ - index used when all releases from Reservoirs 3-6 are constrained.

SPACE

LLJ - index for making releases when all Space allocations are constrained.

(6) C SPACE

LSP - output loop count

SPACE

M - twelve times the number of years in the trace.

MI[illegible] - number of copies of output desired

ML - index used when all releases from Reservoirs 3-6 are constrained.

SPACE

MONTH - index of month number, from 1 to end of trace.

NDRY - sum of the numbers of the reservoir numbers in which current contents exceed the constraints.

C SPACE

NOTDRY - number of reservoirs in which current contents exceed the constraints.

C SPACE

NOTUSE - array of reservoirs removed from the allocation process as the space rule encounters constraints.

(6) SPACE

NRES - number of reservoirs in the system

C

NSIM - number of simulation traces for each system design (set of sizes, etc.) i.e., run.

C

NYR - number of years in each simulation trace

C

PCAP - power capacity of turbines, 1-9, megawatts.

(9) C

PNSTOK - penstock capacity, for Reservoirs 1 or 2, 10^2 ac ft.

C

PNSTK2 - unused water capacity of the penstocks, which limits supplementary release for power from Reservoir 1 or 2, 10^2 ac ft.

SUPPOW

POWER - subroutine of the Monthly Operating Package

PPREC - contents of reservoirs 1-6, corresponding to the current constraint levels, 10^2 ac ft.

(6) C SUPPOW

PREC - contents, Reservoirs 1-6, above current constraint level, to meet target output for energy, 10^2 ac ft.

(6) C SUPPOW

Q - inflow to Reservoirs 1-6, 10^2 ac ft.

(6) C

QB - flow (unregulated) at Bethlehem, 10^2 ac ft.

C

QBREG - regulated flow at Bethlehem, defined as the sum of all reservoir releases plus intermediate flow between the reservoirs and Bethlehem, 10^2 ac ft.

C DIVERT

QDEFCT - the deficit in flow at Bethlehem, defined as the target output for water supply at Bethlehem less the regulated flow at Bethlehem, 10^2 ac ft. If negative, set to zero.

C DIVERT

QDIV - quantity of water available for diversion from the Lehigh to Reservoir 3, 10^2 ac ft., subsequently limited by capacity of the diversion channel and available space in Reservoir 3.

C DIVERT

QDIV2 - volume of second and third supplementary releases, Reservoir 1 or 2, that is diverted to Reservoir 3, 10^2 ac ft.

C

QEXCES - nonnegative difference between flow at Bethlehem and the water supply target, 10^2 ac ft.

C DIVERT

QINT - unregulated inflow for the reach of the river between the reservoirs and Bethlehem, 10^2 ac ft.

C DIVERT

QMIN - minimum flow required below reservoirs 1-6, 10^2 ac ft.

(6) C

QPEN - release available for transmission through the penstock, 10^2 ac ft.

C POWER

QQPEN - release actually made through the penstock, limited by the penstock capacity, 10^2 ac ft.

C POWER

QREC - water available in Reservoirs 1-6 above current constraint levels to meet water supply target at Bethlehem, 10^2 ac ft.

(6) C WATER

QT - flow at Tannery, 10^2 ac ft.

C

QTARG - water supply target at Bethlehem, 10^2 ac ft.

C

REC - amount of water available above constraint levels in Reservoirs 3-6, for allocation of releases in the Space Rule, 10^2 ac ft.

(6) C SPACE

RECADJ - levels in Reservoirs 1-6 corresponding to the second recreation constraint, ft.

(6) C

RECHED - levels in Reservoirs 1-6 corresponding to the first recreation constraint, ft.

(6) C

RECST2 - contents, corresponding to the second recreation constraint level, Reservoirs 1-6, 10^2 ac ft.

(6) C

RECSTO - contents, corresponding to the first recreation constraint level, Reservoirs 1-6, 10^2 ac ft.

(6) C

REL - the release made when only one reservoir remains for allocation by the SPACE rule, 10^2 ac ft.

C SPACE

REQU - total release required from Reservoirs 3-6, by the space rule, to meet target outputs for energy or water supply, 10^2 ac ft.

C WATER or SUPPOW

S - initial contents, reservoirs 1-6, 10^2 ac ft.

(6) C BASIC

SAVG - mean storage, 10^2 ac ft., for Reservoirs 1-6. This value is used to compute the average head and thus the water flow in the turbines.

(6) C POWER

SDEFCT - original value of the power deficit, used to converge on an appropriate value of the required release, mwh.

C SUPPOW

SHORT - release needed to maintain required minimum flows after inevitable releases are made, 10^2 ac ft.

(6) C BASIC

SLOP - volume of releases, Reservoirs 1 or 2, in excess of penstock capacity, 10^2 ac ft.

C

SNEW - current values of contents, Reservoirs 1-6, during operation of the monthly procedure, 10^2 ac ft. At end of monthly procedure, it is final reservoir contents.

(6) C BASIC

SNEWW - current contents of those reservoirs from which allocations are yet to be made, 10^2 ac ft.

(6) C SPACE

SPACE - subroutine of the Monthly Operating Package

C

SQ - intermediate tolerance quantity for checking allocations of releases made in Space rule, 10^2 ac ft.

SPACE

SR - supplementary releases 1-3 from reservoirs 1-6, 10^2 ac ft.

SR_1 - release to meet minimum channel flows

SR_2 - release to meet target output for water supply at Bethlehem

SR_3 - release to meet target output for energy

(3, 6) C BASIC, WATER, SUPPOW

SREQU - release required by Space rule, to be met by releases from Reservoirs 3-6, 10^2 ac ft.

C SPACE

SRI4 - first supplementary release from Reservoir 4

C POWER

SR3K - temporary storage of third supplementary release, Reservoir 1 or 2, 10^2 ac ft.

SUPPOW

SS(6) - original values of contents above the current constraint levels, used in SPACE to maintain a check on the total releases, 10^2 ac ft.

C SPACE

SSPACE - sum of actual releases (ACTUAL) Reservoir 3-6, made by the space rule, 10^2 ac ft.

C SPACE

SSS(6) - original value of storage, reservoirs 1-6, for checking on releases made by SPACE, 10^2 ac ft.

C SPACE

STOL - one percent of REQU - allowable deviation of release allocations in Space rule from REQU, 10^2 ac ft.

SPACE

SUMCAP - sum of total reservoir capacities of those reservoirs from which allocations remain to be made by the space rule, 10^2 ac ft.

C SPACE

SUMNEW - sum of current contents of those reservoirs from which allocations remain to be made by the space rule, 10^2 ac ft.

C SPACE

SUMREC - sum of the amounts of water available above constraint levels in Reservoirs 3-6, for allocation by the space rule, 10^2 ac ft.

C SPACE

SUPPOW - subroutine in Monthly Operating Package

C

TI - fraction of the month during which reservoir contents are at or above dead storage

C POWER

TURBQ - water capacity of turbines 1-9, 10^2 ac ft.

(9) C POWER

TURBR - unutilized water capacity of turbines 1-9, 10^2 ac ft.

(9) C

WATER - subroutine in Monthly Operating Package

C

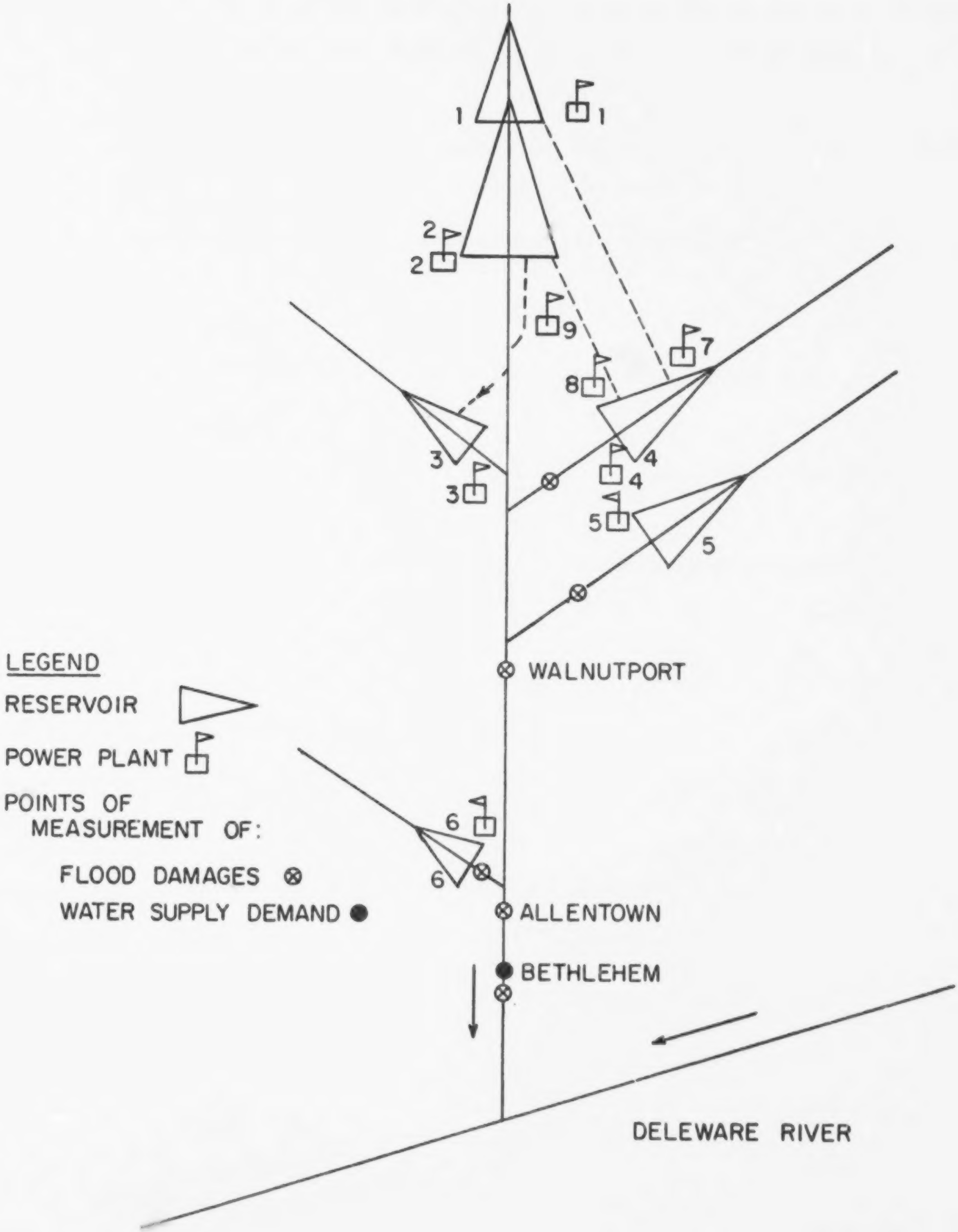

1
1
2
2
9
7
8
3
4
4
3
5
5
WALNUTPORT
LEGEND
RESERVOIR
POWER PLANT
POINTS OF
MEASUREMENT OF:
FLOOD DAMAGES ⊗
WATER SUPPLY DEMAND ●
6
6
ALLENTOWN
BETHLEHEM
DELEWARE RIVER

FIG. 1 THE LEHIGH RIVER SYSTEM

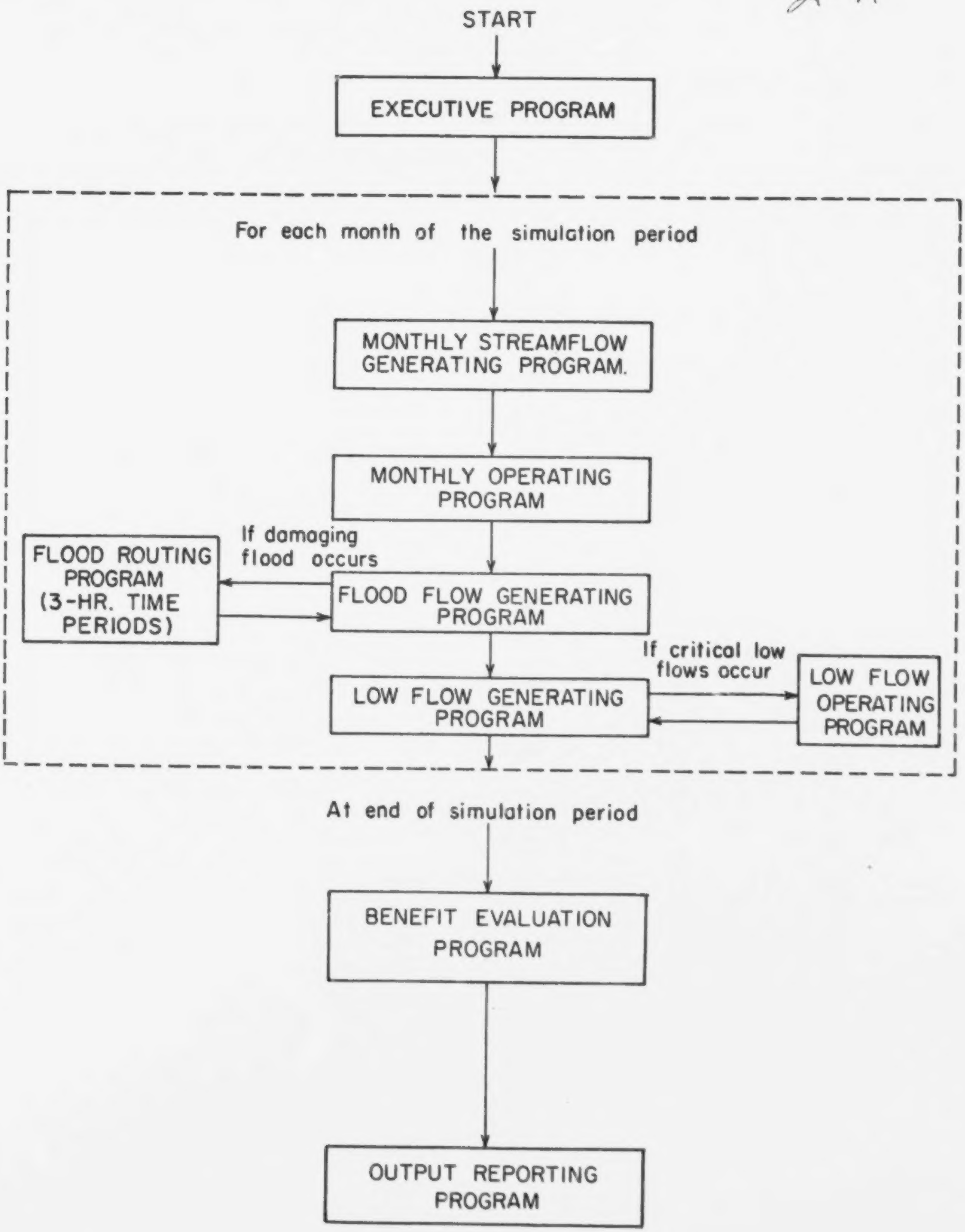
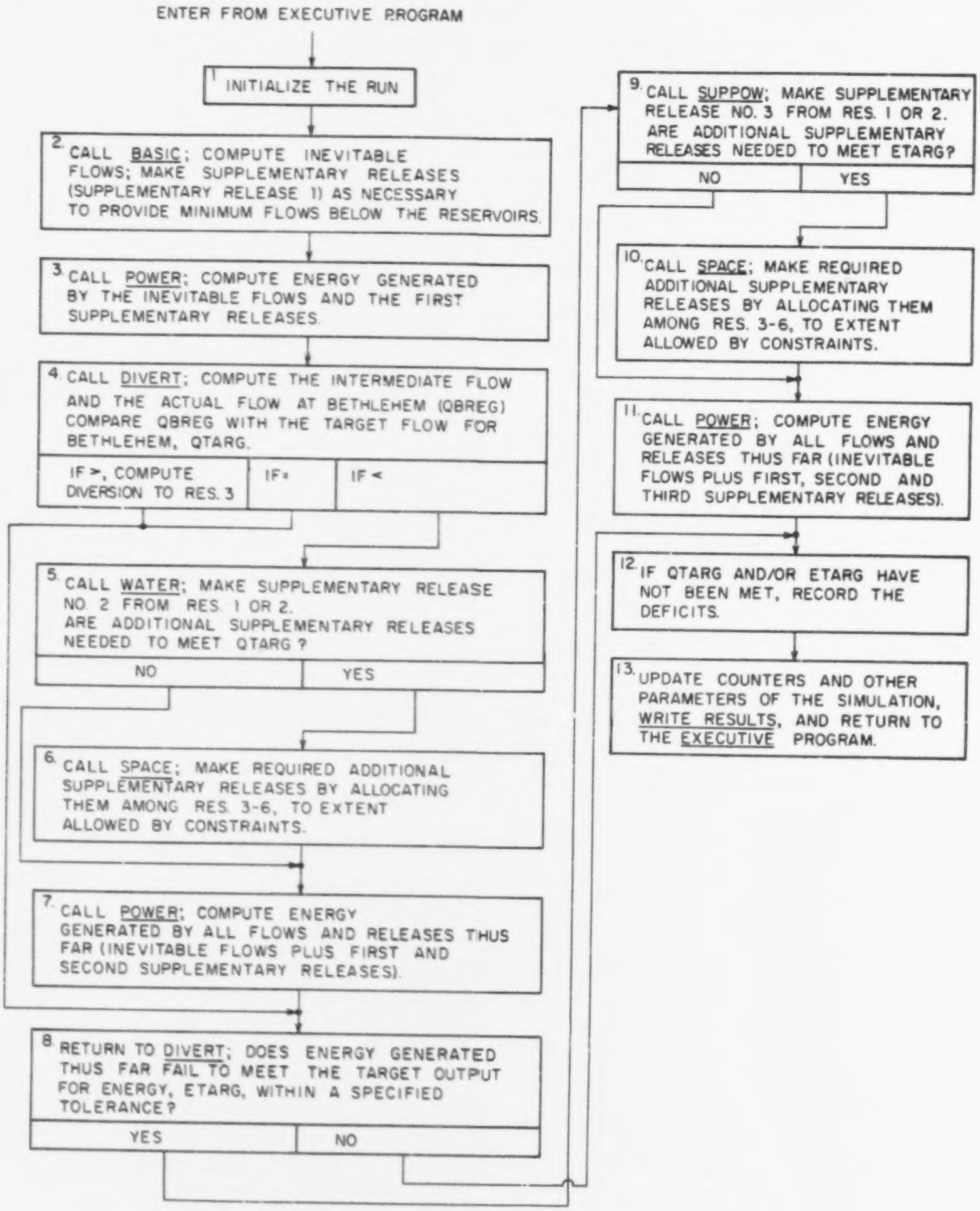
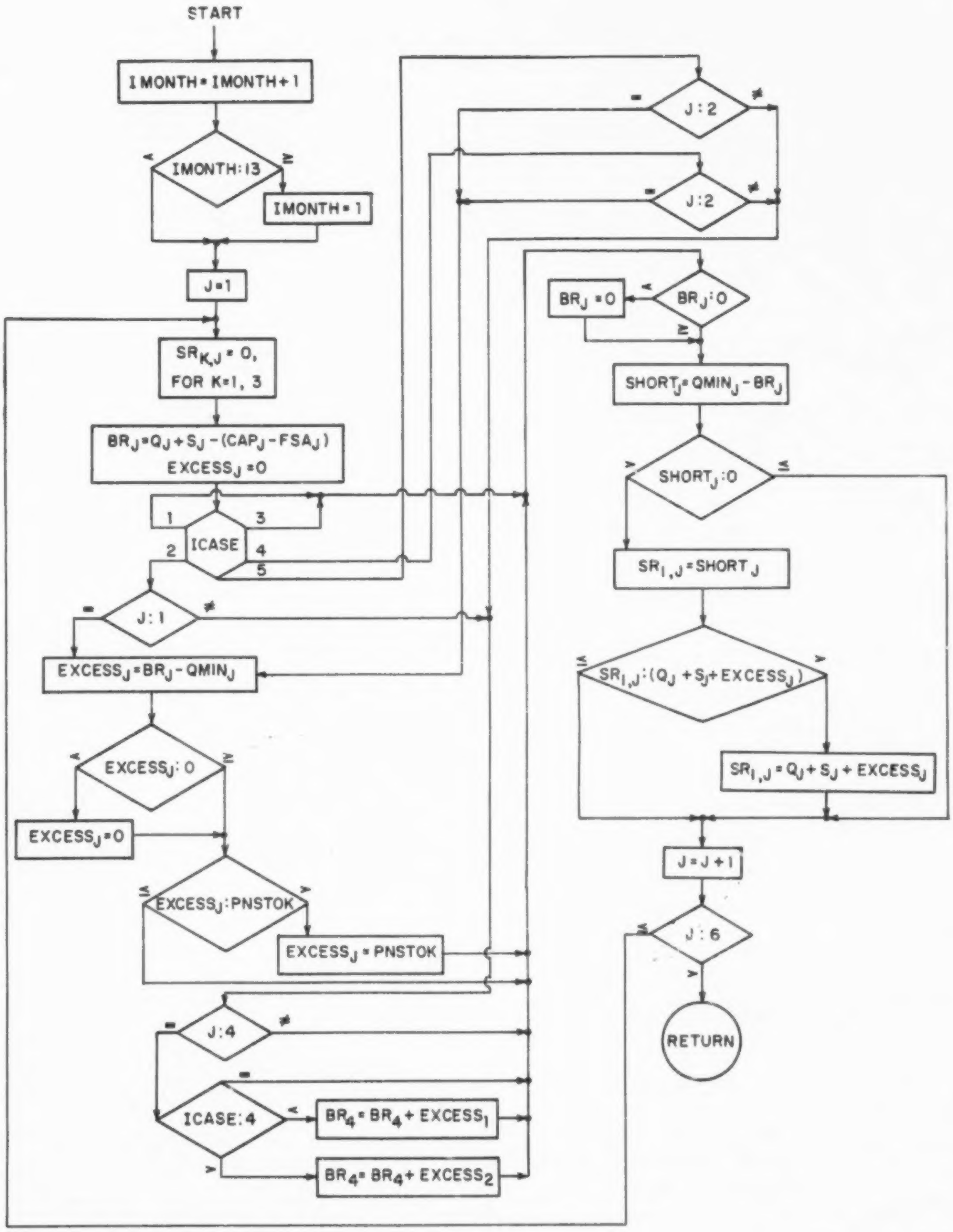

START
EXECUTIVE PROGRAM
For each month of the simulation period
MONTHLY STREAMFLOW GENERATING PROGRAM.
MONTHLY OPERATING PROGRAM
If damaging flood occurs
FLOOD ROUTING PROGRAM (3-HR. TIME PERIODS)
FLOOD FLOW GENERATING PROGRAM
If critical low flows occur
LOW FLOW GENERATING PROGRAM
LOW FLOW OPERATING PROGRAM
At end of simulation period
BENEFIT EVALUATION PROGRAM
OUTPUT REPORTING PROGRAM

FIG. 2 GENERALIZED DIAGRAM OF OVERALL SIMULATION PROGRAM FOR THE LEHIGH RIVER SYSTEM


ENTER FROM EXECUTIVE PROGRAM
1 INITIALIZE THE RUN
2. CALL BASIC; COMPUTE INEVITABLE FLOWS; MAKE SUPPLEMENTARY RELEASES (SUPPLEMENTARY RELEASE 1) AS NECESSARY TO PROVIDE MINIMUM FLOWS BELOW THE RESERVOIRS.
3. CALL POWER; COMPUTE ENERGY GENERATED BY THE INEVITABLE FLOWS AND THE FIRST SUPPLEMENTARY RELEASES.
4. CALL DIVERT; COMPUTE THE INTERMEDIATE FLOW AND THE ACTUAL FLOW AT BETHLEHEM (QBREG) COMPARE QBREG WITH THE TARGET FLOW FOR BETHLEHEM, QTARG.
IF >, COMPUTE DIVERSION TO RES. 3
IF =
IF <
5. CALL WATER; MAKE SUPPLEMENTARY RELEASE NO. 2 FROM RES. 1 OR 2. ARE ADDITIONAL SUPPLEMENTARY RELEASES NEEDED TO MEET QTARG?
NO
YES
6. CALL SPACE; MAKE REQUIRED ADDITIONAL SUPPLEMENTARY RELEASES BY ALLOCATING THEM AMONG RES. 3-6, TO EXTENT ALLOWED BY CONSTRAINTS.
7. CALL POWER; COMPUTE ENERGY GENERATED BY ALL FLOWS AND RELEASES THUS FAR (INEVITABLE FLOWS PLUS FIRST AND SECOND SUPPLEMENTARY RELEASES).
8. RETURN TO DIVERT; DOES ENERGY GENERATED THUS FAR FAIL TO MEET THE TARGET OUTPUT FOR ENERGY, ETARG, WITHIN A SPECIFIED TOLERANCE?
YES
NO
9. CALL SUPPOW; MAKE SUPPLEMENTARY RELEASE NO. 3 FROM RES. 1 OR 2. ARE ADDITIONAL SUPPLEMENTARY RELEASES NEEDED TO MEET ETARG?
NO
YES
10. CALL SPACE; MAKE REQUIRED ADDITIONAL SUPPLEMENTARY RELEASES BY ALLOCATING THEM AMONG RES. 3-6, TO EXTENT ALLOWED BY CONSTRAINTS.
11. CALL POWER; COMPUTE ENERGY GENERATED BY ALL FLOWS AND RELEASES THUS FAR (INEVITABLE FLOWS PLUS FIRST, SECOND AND THIRD SUPPLEMENTARY RELEASES).
12. IF QTARG AND/OR ETARG HAVE NOT BEEN MET, RECORD THE DEFICITS.
13. UPDATE COUNTERS AND OTHER PARAMETERS OF THE SIMULATION, WRITE RESULTS, AND RETURN TO THE EXECUTIVE PROGRAM.

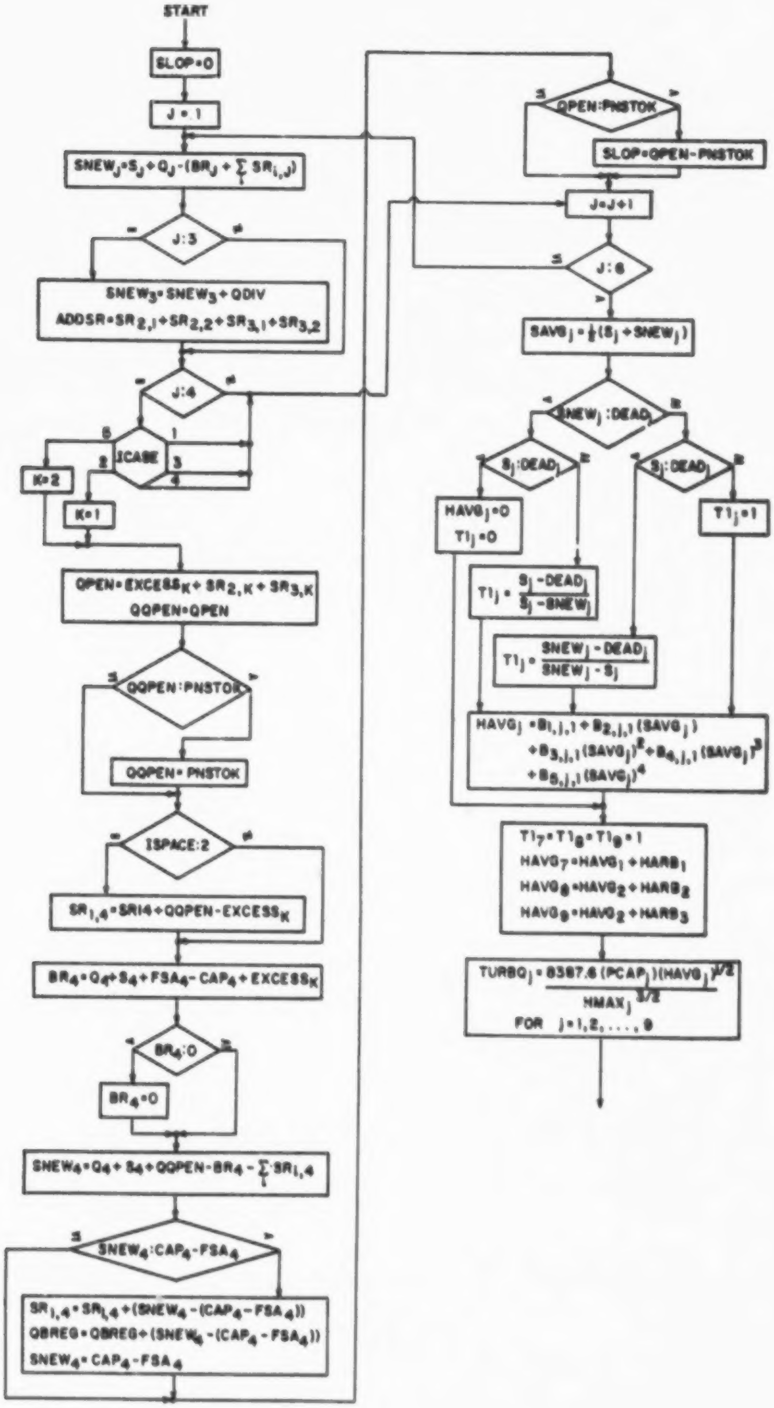
OVERALL PROGRAM FOR THE MONTHLY OPERATING PROCEDURE

Fig. 3

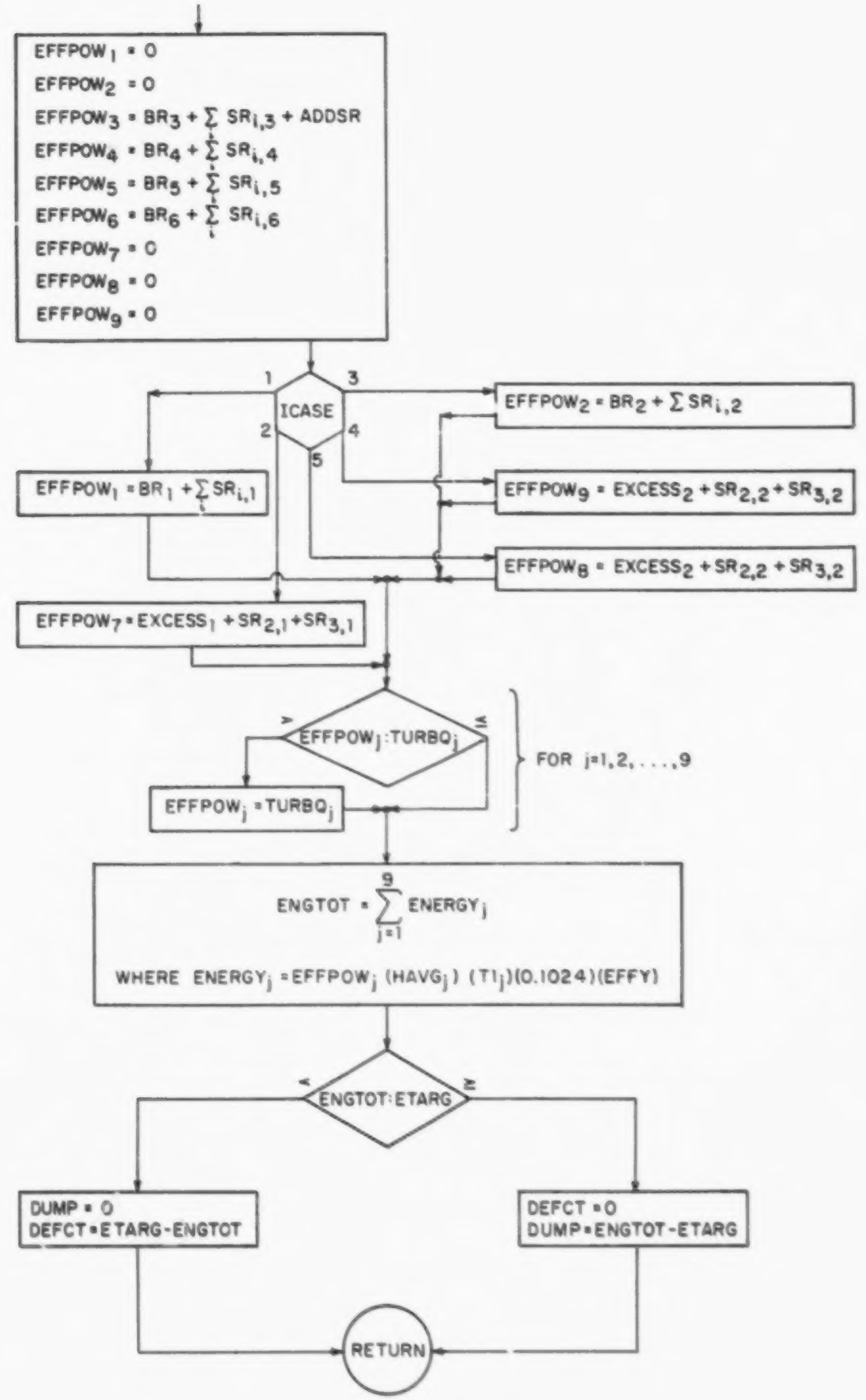

START
IMONTH = IMONTH + 1
IMONTH : 13
<
≥
IMONTH = 1
J = 1
$SR_{K,J} = 0$, FOR K = 1, 3
$BR_J = Q_J + S_J - (CAP_J - FSA_J)$
$EXCESS_J = 0$
ICASE
1
2
3
4
5
J : 1
=
≠
$EXCESS_J = BR_J - QMIN_J$
$EXCESS_J : 0$
<
≥
$EXCESS_J = 0$
$EXCESS_J : PNSTOK$
≤
>
$EXCESS_J = PNSTOK$
J : 4
=
≠
ICASE : 4
=
<
>
$BR_4 = BR_4 + EXCESS_1$
$BR_4 = BR_4 + EXCESS_2$
J : 2
=
≠
J : 2
=
≠
$BR_J : 0$
<
≥
$BR_J = 0$
$SHORT_J = QMIN_J - BR_J$
$SHORT_J : 0$
>
≤
$SR_{1,J} = SHORT_J$
$SR_{1,J} : (Q_J + S_J + EXCESS_J)$
≤
>
$SR_{1,J} = Q_J + S_J + EXCESS_J$
J = J + 1
J : 6
≤
>
RETURN

SUBROUTINE BASIC
Fig. 4

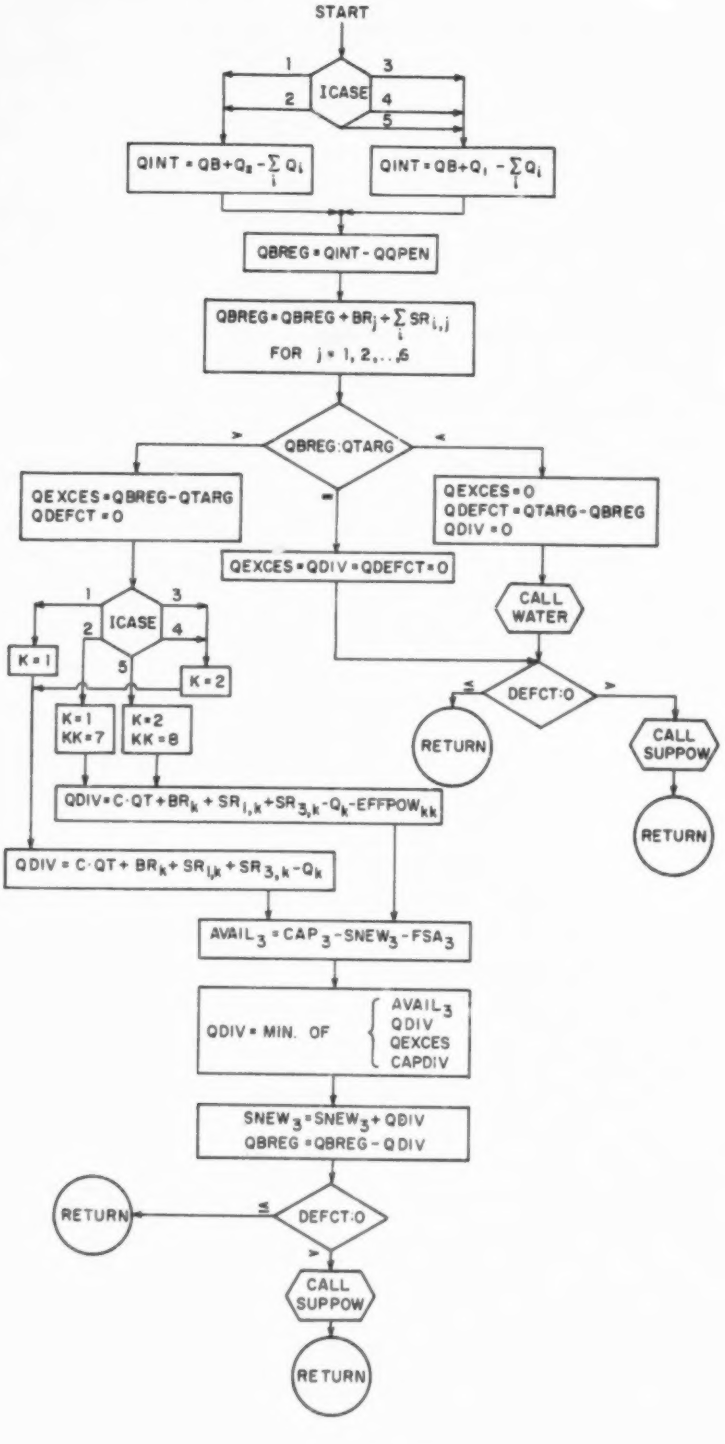
2-44


START
SLOP=0
J=1
$SNEW_J=S_J+Q_J-(BR_J+\sum_i SR_{i,J})$
J:3
=
≠
$SNEW_3=SNEW_3+QDIV$
$ADDSR=SR_{2,1}+SR_{2,2}+SR_{3,1}+SR_{3,2}$
J:4
=
≠
ICASE
5
1
2
3
4
K=2
K=1
$QPEN=EXCESS_K+SR_{2,K}+SR_{3,K}$
QQPEN=QPEN
QQPEN:PNSTOK
≤
>
QQPEN=PNSTOK
ISPACE:2
=
≠
$SR_{1,4}=SR14+QQPEN-EXCESS_K$
$BR_4=Q_4+S_4+FSA_4-CAP_4+EXCESS_K$
$BR_4:0$
<
≥
$BR_4=0$
$SNEW_4=Q_4+S_4+QQPEN-BR_4-\sum_i SR_{i,4}$
$SNEW_4:CAP_4-FSA_4$
≤
>
$SR_{1,4}=SR_{1,4}+(SNEW_4-(CAP_4-FSA_4))$
$QBREG=QBREG+(SNEW_4-(CAP_4-FSA_4))$
$SNEW_4=CAP_4-FSA_4$
QPEN:PNSTOK
≤
>
SLOP=QPEN-PNSTOK
J=J+1
J:6
≤
>
$SAVG_j=\frac{1}{2}(S_j+SNEW_j)$
$SNEW_j:DEAD_j$
<
≥
$S_j:DEAD_j$
<
≥
$S_j:DEAD_j$
<
≥
$HAVG_j=0$
$T1_j=0$
$T1_j=\frac{S_j-DEAD_j}{S_j-SNEW_j}$
$T1_j=\frac{SNEW_j-DEAD_j}{SNEW_j-S_j}$
$T1_j=1$
$HAVG_j=B_{1,j,1}+B_{2,j,1}(SAVG_j)+B_{3,j,1}(SAVG_j)^2+B_{4,j,1}(SAVG_j)^3+B_{5,j,1}(SAVG_j)^4$
$T1_7=T1_8=T1_9=1$
$HAVG_7=HAVG_1+HARB_1$
$HAVG_8=HAVG_2+HARB_2$
$HAVG_9=HAVG_2+HARB_3$
$TURBQ_j=\frac{8387.6(PCAP_j)(HAVG_j)^{1/2}}{HMAX_j^{3/2}}$
FOR j=1,2,...,9

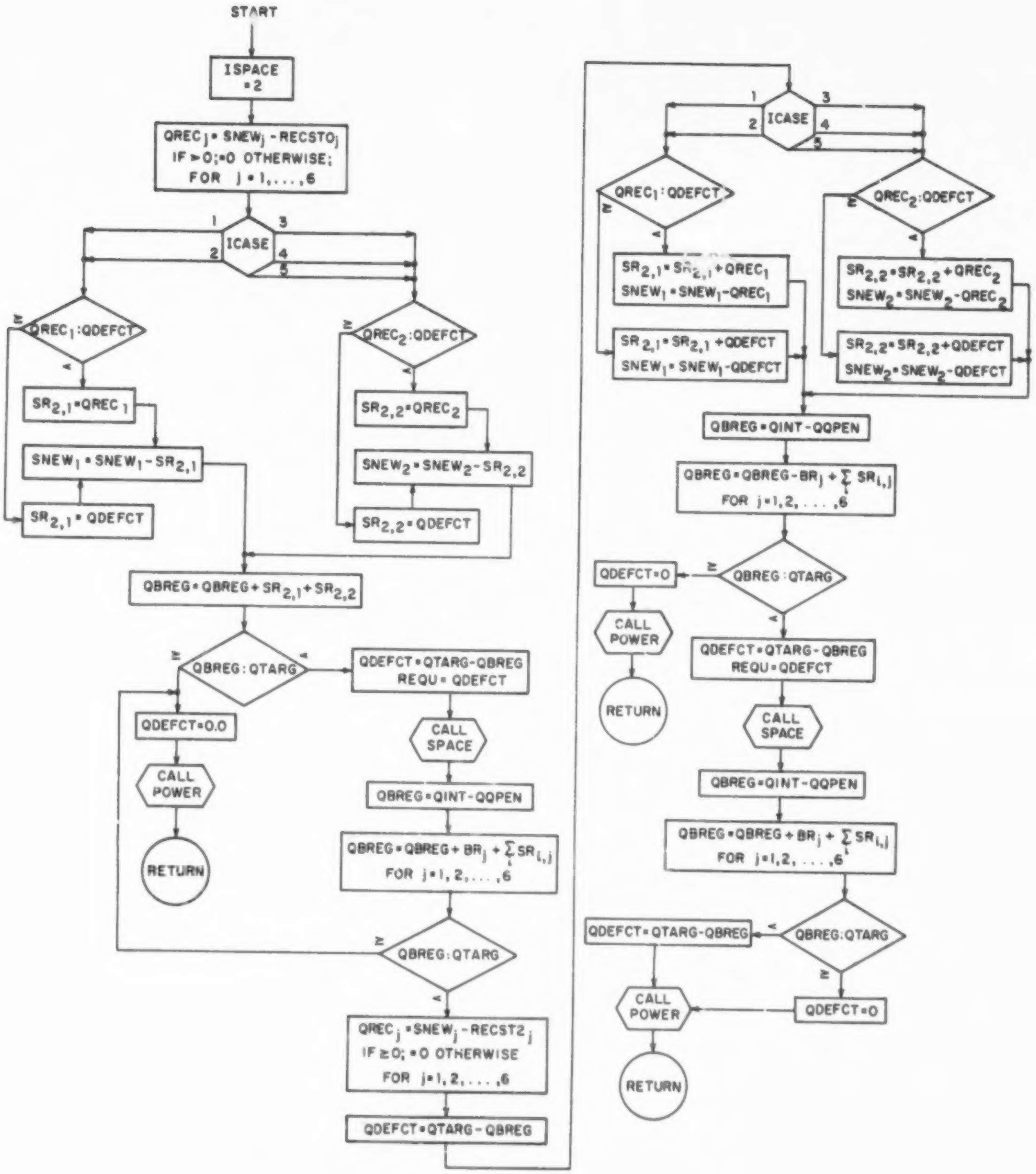
SUBROUTINE POWER
Fig. 5


$EFFPOW_1 = 0$
$EFFPOW_2 = 0$
$EFFPOW_3 = BR_3 + \sum_i SR_{i,3} + ADDSR$
$EFFPOW_4 = BR_4 + \sum_i SR_{i,4}$
$EFFPOW_5 = BR_5 + \sum_i SR_{i,5}$
$EFFPOW_6 = BR_6 + \sum_i SR_{i,6}$
$EFFPOW_7 = 0$
$EFFPOW_8 = 0$
$EFFPOW_9 = 0$
ICASE
1
2
3
4
5
$EFFPOW_2 = BR_2 + \sum SR_{i,2}$
$EFFPOW_1 = BR_1 + \sum_i SR_{i,1}$
$EFFPOW_9 = EXCESS_2 + SR_{2,2} + SR_{3,2}$
$EFFPOW_8 = EXCESS_2 + SR_{2,2} + SR_{3,2}$
$EFFPOW_7 = EXCESS_1 + SR_{2,1} + SR_{3,1}$
$>$
$EFFPOW_j : TURBQ_j$
$\leq$
$EFFPOW_j = TURBQ_j$
FOR $j = 1, 2, \ldots, 9$
$ENGTOT = \sum_{j=1}^{9} ENERGY_j$
WHERE $ENERGY_j = EFFPOW_j (HAVG_j)(T1_j)(0.1024)(EFFY)$
$<$
ENGTOT : ETARG
$\geq$
DUMP = 0
DEFCT = ETARG - ENGTOT
DEFCT = 0
DUMP = ENGTOT - ETARG
RETURN

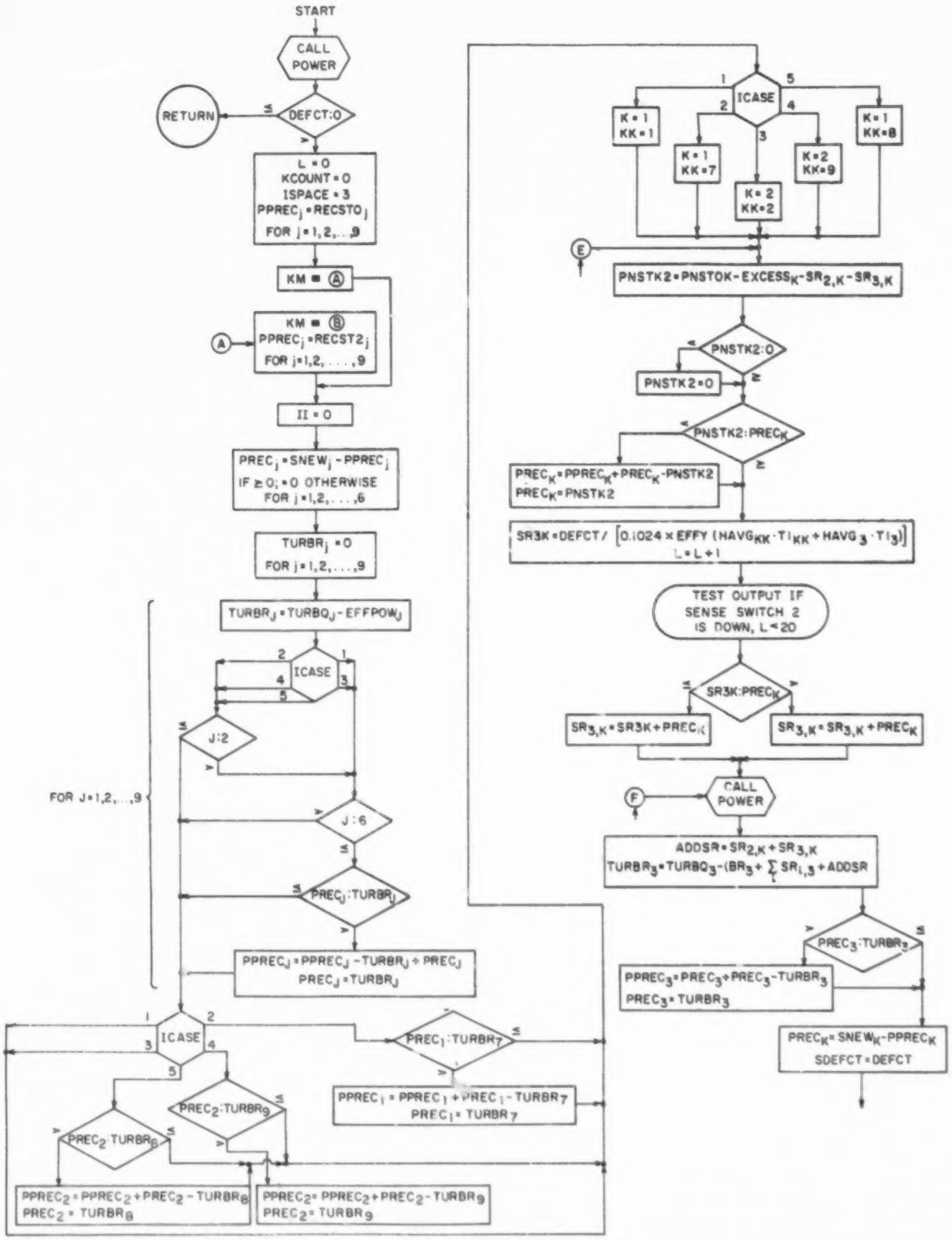
SUBROUTINE POWER (Cont.)
Fig. 5


START
ICASE
1
2
3
4
5
QINT = QB + Q_2 − $\sum_i Q_i$
QINT = QB + Q_1 − $\sum_i Q_i$
QBREG = QINT − QQPEN
QBREG = QBREG + BR_j + $\sum_i SR_{i,j}$
FOR j = 1, 2, ... ,6
QBREG:QTARG
>
<
=
QEXCES = QBREG − QTARG
QDEFCT = 0
QEXCES = QDIV = QDEFCT = 0
QEXCES = 0
QDEFCT = QTARG − QBREG
QDIV = 0
CALL WATER
ICASE
1
2
3
4
5
K = 1
K = 2
K = 1
KK = 7
K = 2
KK = 8
DEFCT:0
≤
>
RETURN
CALL SUPPOW
RETURN
QDIV = C·QT + BR_k + $SR_{1,k}$ + $SR_{3,k}$ − Q_k − $EFFPOW_{kk}$
QDIV = C·QT + BR_k + $SR_{1,k}$ + $SR_{3,k}$ − Q_k
$AVAIL_3$ = CAP_3 − $SNEW_3$ − FSA_3
QDIV = MIN. OF
$AVAIL_3$
QDIV
QEXCES
CAPDIV
$SNEW_3$ = $SNEW_3$ + QDIV
QBREG = QBREG − QDIV
DEFCT:0
≤
>
RETURN
CALL SUPPOW
RETURN

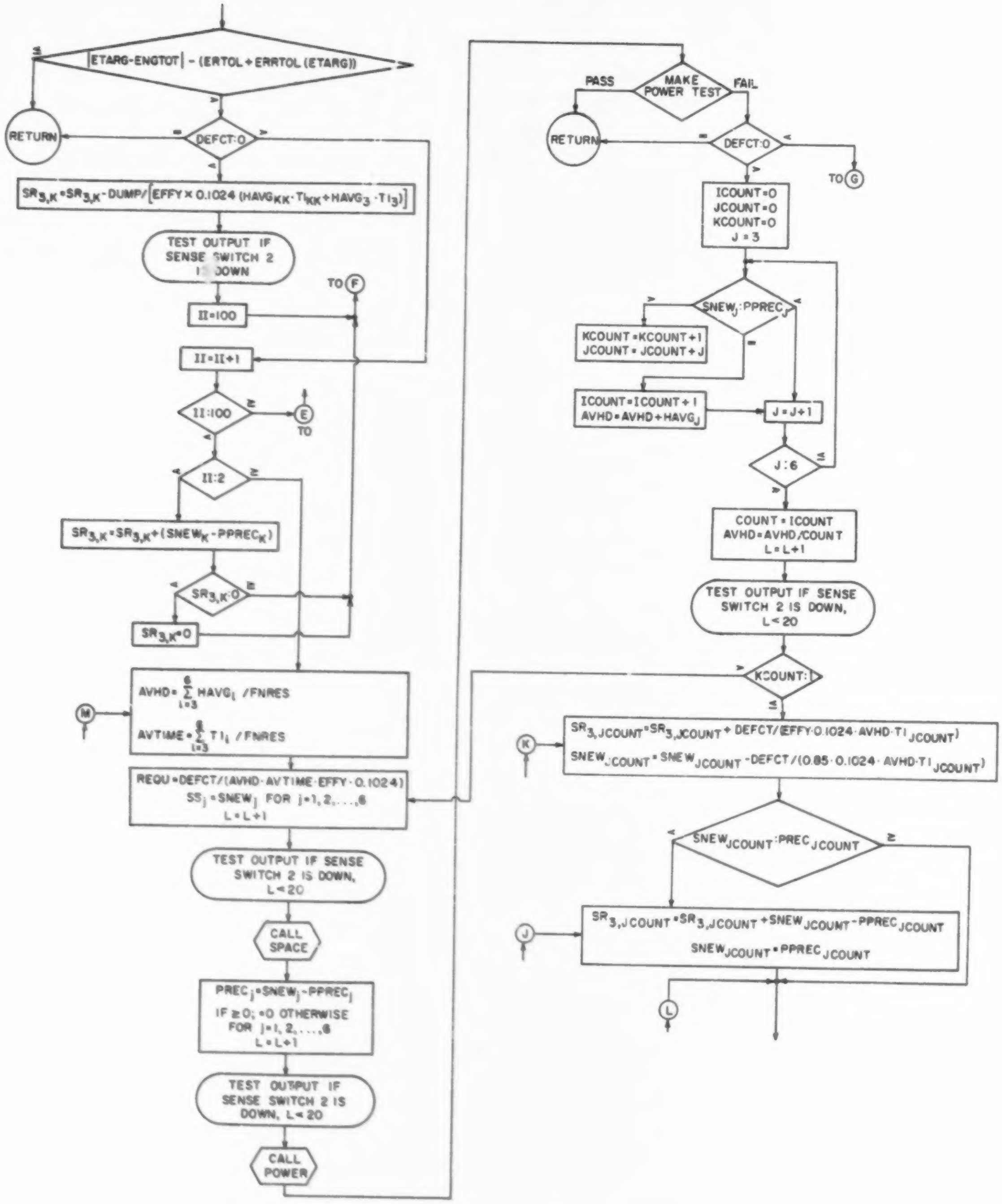
SUBROUTINE DIVERT
Fig. 6


START
ISPACE =2
QREC$_j$ = SNEW$_j$ - RECSTO$_j$ IF > 0; =0 OTHERWISE; FOR j = 1, . . . ,6
ICASE
1
2
3
4
5
≥
QREC$_1$:QDEFCT
<
SR$_{2,1}$=QREC$_1$
SNEW$_1$ = SNEW$_1$ - SR$_{2,1}$
SR$_{2,1}$ = QDEFCT
≥
QREC$_2$:QDEFCT
<
SR$_{2,2}$=QREC$_2$
SNEW$_2$ = SNEW$_2$ - SR$_{2,2}$
SR$_{2,2}$ = QDEFCT
QBREG = QBREG + SR$_{2,1}$ + SR$_{2,2}$
≥
QBREG : QTARG
<
QDEFCT=0.0
CALL POWER
RETURN
QDEFCT = QTARG-QBREG REQU = QDEFCT
CALL SPACE
QBREG = QINT - QQPEN
QBREG = QBREG + BR$_j$ + $\sum_i$ SR$_{i,j}$ FOR j = 1, 2, . . . ,6
≥
QBREG:QTARG
<
QREC$_j$ = SNEW$_j$ - RECST2$_j$ IF ≥ 0; = 0 OTHERWISE FOR j = 1, 2, . . . ,6
QDEFCT = QTARG - QBREG
ICASE
1
2
3
4
5
QREC$_1$: QDEFCT
≥
<
SR$_{2,1}$ = SR$_{2,1}$ + QREC$_1$
SNEW$_1$ = SNEW$_1$ - QREC$_1$
SR$_{2,1}$ = SR$_{2,1}$ + QDEFCT
SNEW$_1$ = SNEW$_1$ - QDEFCT
≥
QREC$_2$:QDEFCT
<
SR$_{2,2}$ = SR$_{2,2}$ + QREC$_2$
SNEW$_2$ = SNEW$_2$ - QREC$_2$
SR$_{2,2}$ = SR$_{2,2}$ + QDEFCT
SNEW$_2$ = SNEW$_2$ - QDEFCT
QBREG = QINT - QQPEN
QBREG = QBREG - BR$_j$ + $\sum_i$ SR$_{i,j}$ FOR j = 1, 2, . . . ,6
QDEFCT=0
≥
QBREG : QTARG
<
CALL POWER
RETURN
QDEFCT = QTARG - QBREG REQU = QDEFCT
CALL SPACE
QBREG = QINT - QQPEN
QBREG = QBREG + BR$_j$ + $\sum_i$ SR$_{i,j}$ FOR j = 1, 2, . . . ,6
QDEFCT = QTARG-QBREG
<
QBREG:QTARG
≥
CALL POWER
QDEFCT=0
RETURN

SUBROUTINE WATER
Fig. 7

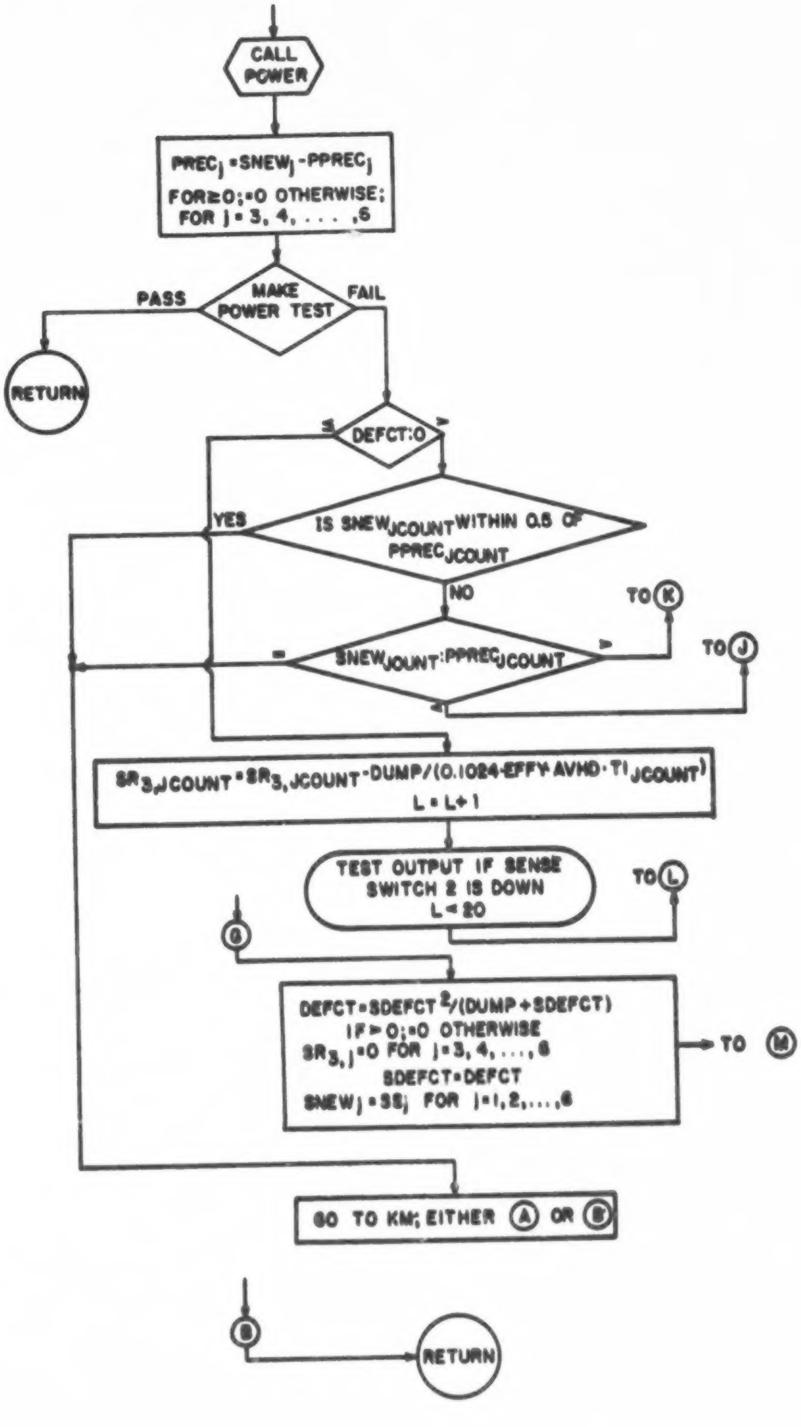

START
CALL POWER
RETURN
DEFCT:0
≤
>
L = 0
KCOUNT = 0
ISPACE = 3
PPREC$_j$ = RECSTO$_j$
FOR j = 1,2,...,9
KM = Ⓐ
KM = Ⓑ
PPREC$_j$ = RECST2$_j$
FOR j = 1,2,...,9
Ⓐ
II = 0
PREC$_j$ = SNEW$_j$ - PPREC$_j$
IF ≥ 0; = 0 OTHERWISE
FOR j = 1,2,...,6
TURBR$_j$ = 0
FOR j = 1,2,...,9
TURBR$_J$ = TURBQ$_J$ - EFFPOW$_J$
FOR J = 1,2,...,9
ICASE
1
2
3
4
5
J:2
J:6
PREC$_J$:TURBR$_J$
PPREC$_J$ = PPREC$_J$ - TURBR$_J$ + PREC$_J$
PREC$_J$ = TURBR$_J$
ICASE
PREC$_1$:TURBR$_7$
PPREC$_1$ = PPREC$_1$ + PREC$_1$ - TURBR$_7$
PREC$_1$ = TURBR$_7$
PREC$_2$:TURBR$_9$
PREC$_2$:TURBR$_8$
PPREC$_2$ = PPREC$_2$ + PREC$_2$ - TURBR$_8$
PREC$_2$ = TURBR$_8$
PPREC$_2$ = PPREC$_2$ + PREC$_2$ - TURBR$_9$
PREC$_2$ = TURBR$_9$
ICASE
K = 1
KK = 1
K = 1
KK = 7
K = 2
KK = 2
K = 2
KK = 9
K = 1
KK = 8
Ⓔ
PNSTK2 = PNSTOK - EXCESS$_K$ - SR$_{2,K}$ - SR$_{3,K}$
PNSTK2:0
PNSTK2 = 0
PNSTK2:PREC$_K$
PREC$_K$ = PPREC$_K$ + PREC$_K$ - PNSTK2
PREC$_K$ = PNSTK2
SR3K = DEFCT / [0.1024 × EFFY (HAVG$_{KK}$ · TI$_{KK}$ + HAVG$_3$ · TI$_3$)]
L = L + 1
TEST OUTPUT IF SENSE SWITCH 2 IS DOWN, L < 20
SR3K:PREC$_K$
SR$_{3,K}$ = SR3K + PREC$_K$
SR$_{3,K}$ = SR$_{3,K}$ + PREC$_K$
Ⓕ
CALL POWER
ADDSR = SR$_{2,K}$ + SR$_{3,K}$
TURBR$_3$ = TURBQ$_3$ - (BR$_3$ + $\sum_i$ SR$_{i,3}$ + ADDSR
PREC$_3$:TURBR$_3$
PPREC$_3$ = PREC$_3$ + PREC$_3$ - TURBR$_3$
PREC$_3$ = TURBR$_3$
PREC$_K$ = SNEW$_K$ - PPREC$_K$
SDEFCT = DEFCT

SUBROUTINE SUPPOW
Fig. 8


|ETARG-ENGTOT| - (ERTOL + ERRTOL (ETARG))
≤
>
RETURN
DEFCT:0
=
>
<
SR3,K = SR3,K - DUMP/[EFFY × 0.1024 (HAVGKK · TIKK + HAVG3 · TI3)]
TEST OUTPUT IF SENSE SWITCH 2 IS DOWN
TO F
II = 100
II = II + 1
II:100
≥
TO E
<
II:2
<
≥
SR3,K = SR3,K + (SNEWK - PPRECK)
SR3,K:0
<
≥
SR3,K = 0
M
AVHD = Σ(i=3 to 6) HAVGi /FNRES
AVTIME = Σ(i=3 to 6) TIi /FNRES
REQU = DEFCT/(AVHD · AVTIME · EFFY · 0.1024)
SSj = SNEWj FOR j = 1, 2, ..., 6
L = L + 1
TEST OUTPUT IF SENSE SWITCH 2 IS DOWN, L < 20
CALL SPACE
PRECj = SNEWj - PPRECj
IF ≥ 0; = 0 OTHERWISE
FOR j = 1, 2, ..., 6
L = L + 1
TEST OUTPUT IF SENSE SWITCH 2 IS DOWN, L < 20
CALL POWER
PASS
MAKE POWER TEST
FAIL
RETURN
DEFCT:0
=
<
>
TO G
ICOUNT = 0
JCOUNT = 0
KCOUNT = 0
J = 3
SNEWJ : PPRECJ
>
<
=
KCOUNT = KCOUNT + 1
JCOUNT = JCOUNT + J
ICOUNT = ICOUNT + 1
AVHD = AVHD + HAVGJ
J = J + 1
J:6
≤
>
COUNT = ICOUNT
AVHD = AVHD/COUNT
L = L + 1
TEST OUTPUT IF SENSE SWITCH 2 IS DOWN, L < 20
KCOUNT:1
>
≤
K
SR3,JCOUNT = SR3,JCOUNT + DEFCT/(EFFY · 0.1024 · AVHD · TIJCOUNT)
SNEWJCOUNT = SNEWJCOUNT - DEFCT/(0.85 · 0.1024 · AVHD · TIJCOUNT)
SNEWJCOUNT : PRECJCOUNT
<
≥
J
SR3,JCOUNT = SR3,JCOUNT + SNEWJCOUNT - PPRECJCOUNT
SNEWJCOUNT = PPRECJCOUNT
L

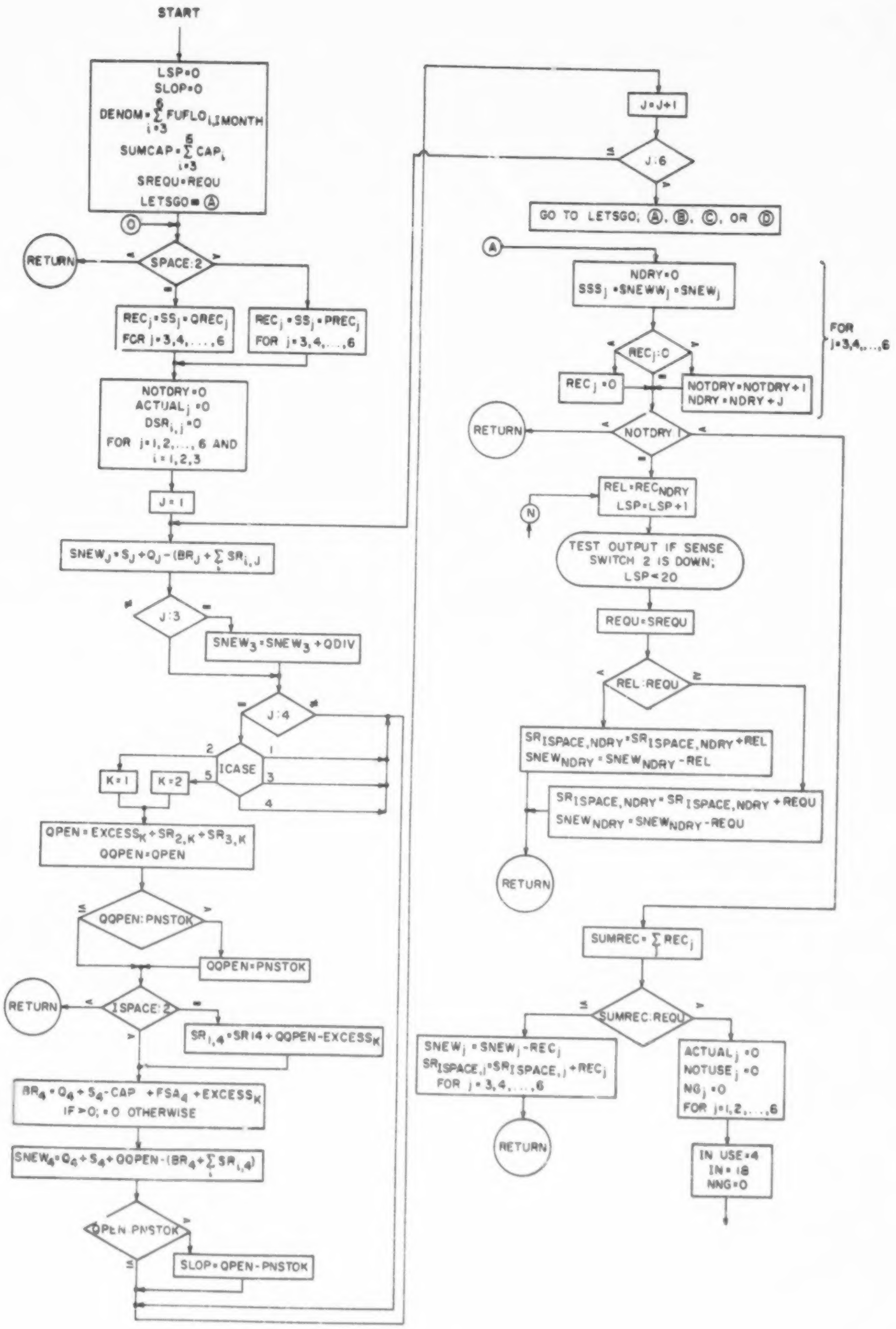
SUBROUTINE SUPPOW (Cont.)
Fig. 8

2-50


CALL POWER
PREC_j = SNEW_j - PPREC_j
FOR ≥ 0; = 0 OTHERWISE;
FOR j = 3, 4, . . . , 6
PASS
MAKE POWER TEST
FAIL
RETURN
≤
DEFCT:0
>
YES
IS SNEW_JCOUNT WITHIN 0.5 OF PPREC_JCOUNT
NO
TO Ⓚ
=
SNEW_JOUNT:PPREC_JCOUNT
>
TO Ⓙ
<
SR_3,JCOUNT = SR_3,JCOUNT - DUMP/(0.1024·EFFY·AVHD·TI_JCOUNT)
L = L + 1
TEST OUTPUT IF SENSE SWITCH 2 IS DOWN
L < 20
TO Ⓛ
Ⓖ
DEFCT = SDEFCT²/(DUMP + SDEFCT)
IF > 0; = 0 OTHERWISE
SR_3,j = 0 FOR j = 3, 4, . . . , 6
SDEFCT = DEFCT
SNEW_j = SS_j FOR j = 1, 2, . . . , 6
TO Ⓜ
GO TO KM; EITHER Ⓐ OR Ⓑ
Ⓑ
RETURN

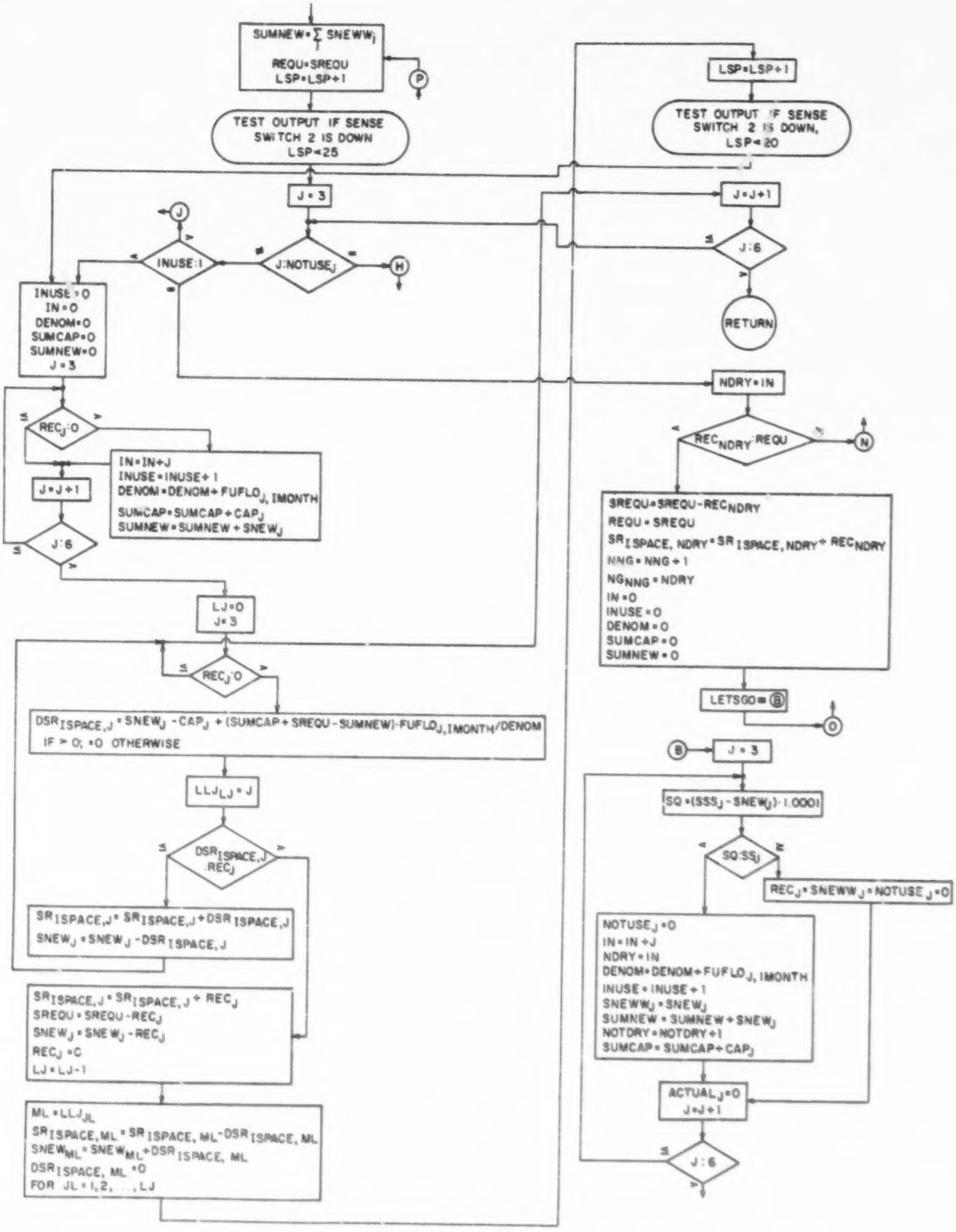
SUBROUTINE SUPPOW (Concluded)

Fig. 8


START
LSP=0
SLOP=0
DENOM=$\sum_{i=3}^{6}$ FUFLO$_{i,IMONTH}$
SUMCAP=$\sum_{i=3}^{6}$ CAP$_i$
SREQU=REQU
LETSGO=Ⓐ
Ⓞ
RETURN
SPACE:2
REC$_j$=SS$_j$=QREC$_j$ FOR j=3,4,...,6
REC$_j$=SS$_j$=PREC$_j$ FOR j=3,4,...,6
NOTDRY=0
ACTUAL$_j$=0
DSR$_{i,j}$=0
FOR j=1,2,...,6 AND i=1,2,3
J=1
SNEW$_J$=S$_J$+Q$_J$-(BR$_J$+$\sum_i$SR$_{i,J}$)
J:3
SNEW$_3$=SNEW$_3$+QDIV
J:4
ICASE
1
2
3
4
5
K=1
K=2
QPEN=EXCESS$_K$+SR$_{2,K}$+SR$_{3,K}$
QOPEN=QPEN
QOPEN:PNSTOK
QOPEN=PNSTOK
RETURN
ISPACE:2
SR$_{1,4}$=SR14+QOPEN-EXCESS$_K$
BR$_4$=Q$_4$+S$_4$-CAP +FSA$_4$+EXCESS$_K$ IF>0; =0 OTHERWISE
SNEW$_4$=Q$_4$+S$_4$+QOPEN-(BR$_4$+$\sum_i$SR$_{i,4}$)
QPEN:PNSTOK
SLOP=QPEN-PNSTOK
J=J+1
J:6
GO TO LETSGO; Ⓐ, Ⓑ, Ⓒ, OR Ⓓ
Ⓐ
NDRY=0
SSS$_j$=SNEWW$_j$=SNEW$_j$
FOR j=3,4,...,6
REC$_j$:0
REC$_j$=0
NOTDRY=NOTDRY+1
NDRY=NDRY+J
RETURN
NOTDRY:1
REL=REC$_{NDRY}$
LSP=LSP+1
Ⓝ
TEST OUTPUT IF SENSE SWITCH 2 IS DOWN; LSP<20
REQU=SREQU
REL:REQU
SR$_{ISPACE,NDRY}$=SR$_{ISPACE,NDRY}$+REL
SNEW$_{NDRY}$=SNEW$_{NDRY}$-REL
SR$_{ISPACE,NDRY}$=SR$_{ISPACE,NDRY}$+REQU
SNEW$_{NDRY}$=SNEW$_{NDRY}$-REQU
RETURN
SUMREC=$\sum_j$REC$_j$
SUMREC:REQU
SNEW$_j$=SNEW$_j$-REC$_j$
SR$_{ISPACE,j}$=SR$_{ISPACE,j}$+REC$_j$
FOR j=3,4,...,6
RETURN
ACTUAL$_j$=0
NOTUSE$_j$=0
NG$_j$=0
FOR j=1,2,...,6
IN USE=4
IN=18
NNG=0

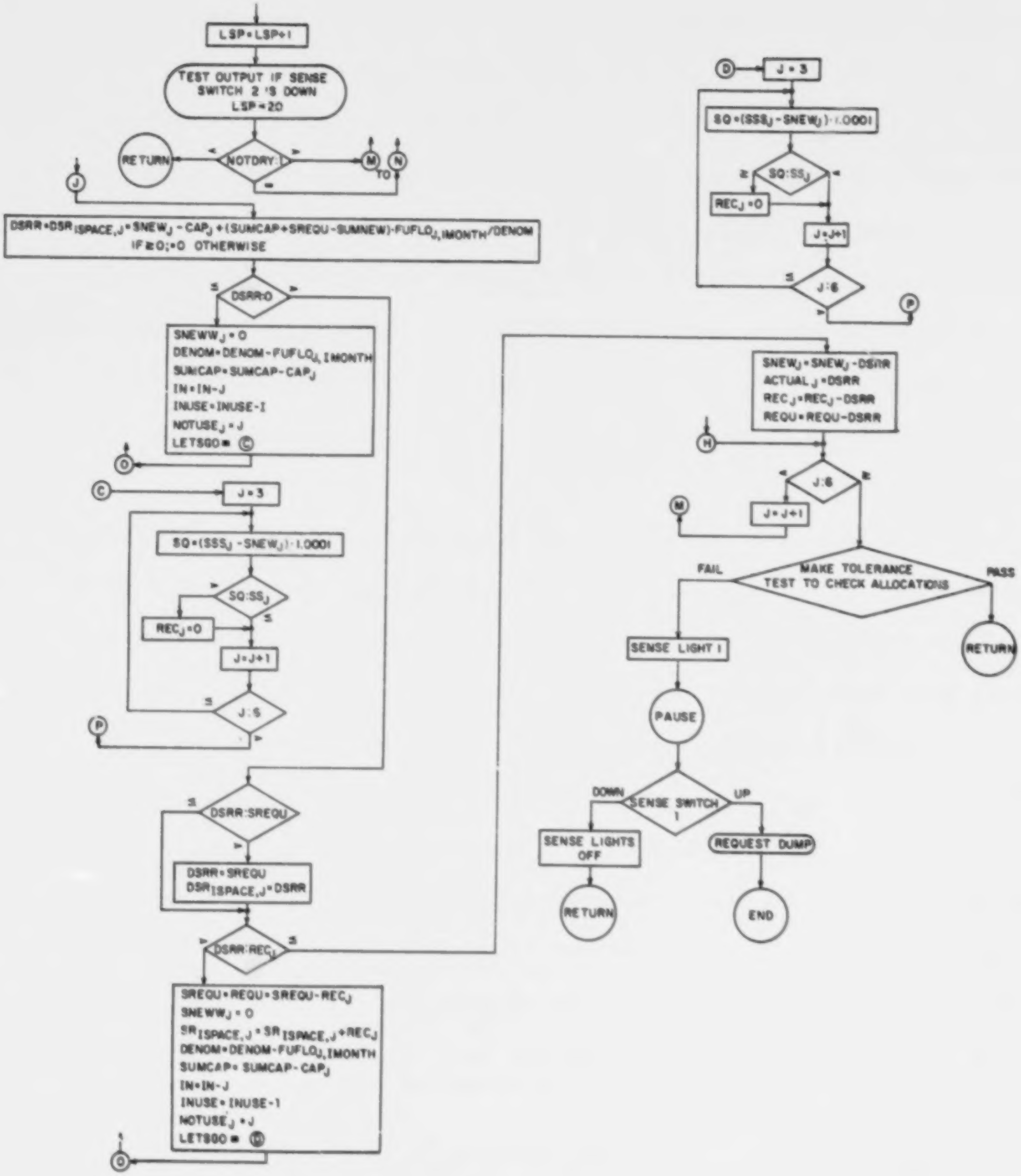
SUBROUTINE SPACE

Fig. 9


$SUMNEW = \sum_J SNEWW_J$
REQU = SREQU
LSP = LSP + 1
P
TEST OUTPUT IF SENSE SWITCH 2 IS DOWN LSP < 25
J = 3
J
INUSE : 1
<
>
=
$J : NOTUSE_J$
≠
=
H
INUSE = 0
IN = 0
DENOM = 0
SUMCAP = 0
SUMNEW = 0
J = 3
$REC_J : 0$
≤
>
IN = IN + J
INUSE = INUSE + 1
$DENOM = DENOM + FUFLO_{J, IMONTH}$
$SUMCAP = SUMCAP + CAP_J$
$SUMNEW = SUMNEW + SNEW_J$
J = J + 1
J : 6
≤
>
LJ = 0
J = 3
$REC_J : 0$
≤
>
$DSR_{ISPACE,J} = SNEW_J - CAP_J + (SUMCAP + SREQU - SUMNEW) \cdot FUFLO_{J,IMONTH}/DENOM$
IF > 0; = 0 OTHERWISE
$LLJ_{LJ} = J$
$DSR_{ISPACE,J} : REC_J$
≤
>
$SR_{ISPACE,J} = SR_{ISPACE,J} + DSR_{ISPACE,J}$
$SNEW_J = SNEW_J - DSR_{ISPACE,J}$
$SR_{ISPACE,J} = SR_{ISPACE,J} + REC_J$
$SREQU = SREQU - REC_J$
$SNEW_J = SNEW_J - REC_J$
$REC_J = C$
LJ = LJ - 1
$ML = LLJ_{JL}$
$SR_{ISPACE,ML} = SR_{ISPACE,ML} - DSR_{ISPACE,ML}$
$SNEW_{ML} = SNEW_{ML} + DSR_{ISPACE,ML}$
$DSR_{ISPACE,ML} = 0$
FOR JL = 1, 2, . . . , LJ
LSP = LSP + 1
TEST OUTPUT IF SENSE SWITCH 2 IS DOWN, LSP < 20
J = J + 1
J : 6
≤
>
RETURN
NDRY = IN
$REC_{NDRY} : REQU$
<
≥
N
$SREQU = SREQU - REC_{NDRY}$
REQU = SREQU
$SR_{ISPACE,NDRY} = SR_{ISPACE,NDRY} + REC_{NDRY}$
NNG = NNG + 1
$NG_{NNG} = NDRY$
IN = 0
INUSE = 0
DENOM = 0
SUMCAP = 0
SUMNEW = 0
LETSGO = B
O
B
J = 3
$SQ = (SSS_J - SNEW_J) \cdot 1.0001$
$SQ : SS_J$
<
≥
$REC_J = SNEWW_J = NOTUSE_J = 0$
$NOTUSE_J = 0$
IN = IN + J
NDRY = IN
$DENOM = DENOM + FUFLO_{J, IMONTH}$
INUSE = INUSE + 1
$SNEWW_J = SNEW_J$
$SUMNEW = SUMNEW + SNEW_J$
NOTDRY = NOTDRY + 1
$SUMCAP = SUMCAP + CAP_J$
$ACTUAL_J = 0$
J = J + 1
J : 6
≤
>

SUBROUTINE SPACE (Cont.)

Fig. 9


LSP=LSP+1
TEST OUTPUT IF SENSE SWITCH 2 IS DOWN LSP=20
RETURN
NOTDRY:1
M
TO
N
J
DSRR=DSR ISPACE,J + SNEW J - CAP J + (SUMCAP+SREQU-SUMNEW)·FUFLO J,IMONTH /DENOM IF ≥ 0; = 0 OTHERWISE
DSRR:0
SNEWW J = 0
DENOM = DENOM - FUFLO J,IMONTH
SUMCAP = SUMCAP - CAP J
IN = IN - J
INUSE = INUSE - 1
NOTUSE J = J
LETSGO = C
O
C
J = 3
SQ = (SSS J - SNEW J)·1.0001
SQ:SS J
REC J = 0
J = J+1
J:5
P
DSRR:SREQU
DSRR = SREQU
DSR ISPACE,J = DSRR
DSRR:REC J
SREQU = REQU = SREQU - REC J
SNEWW J = 0
SR ISPACE,J = SR ISPACE,J + REC J
DENOM = DENOM - FUFLO J,IMONTH
SUMCAP = SUMCAP - CAP J
IN = IN - J
INUSE = INUSE - 1
NOTUSE J = J
LETSGO = D
O
D
J = 3
SQ = (SSS J - SNEW J)·1.0001
SQ:SS J
REC J = 0
J = J+1
J:6
P
SNEW J = SNEW J - DSRR
ACTUAL J = DSRR
REC J = REC J - DSRR
REQU = REQU - DSRR
H
J:6
M
J = J+1
FAIL
MAKE TOLERANCE TEST TO CHECK ALLOCATIONS
PASS
RETURN
SENSE LIGHT 1
PAUSE
DOWN
SENSE SWITCH 1
UP
SENSE LIGHTS OFF
REQUEST DUMP
RETURN
END

SUBROUTINE SPACE (Concluded)
Fig. 9

CHAPTER THREE

Queueing Model for Pollution Transport in Streams

General Problem: Model SQ-SX

This chapter is concerned with the development of mathematical techniques for the formulation of the complex process of pollution transport in streams. Model SQ-SX is a stochastic construct employing the mathematical techniques of queueing theory, probability theory, and finite Markov chains. This particular model is unique, in that it is the first time that an attempt has been made to queue two interdependent stochastic variates simultaneously, namely, the water and a pollutant. This is the first of a series of stochastic models to be studied along these lines and therefore, in some areas, it may appear to be unsophisticated.

The work to date includes the stochastic formulation of: hydrology with lag-one serial correlation between inflows; and pollutants with both lag-one serial correlation and a decay mechanism for the study of degrading organic wastes, thermal wastes, and radioactive wastes.

The objectives of this study are: for any given (1) pattern of flow, (2) temperature, (3) solar radiation, (4) shape of channel and location of dams, (5) pattern and type of loading of pollutants, to predict the quality and quantity of water at given control points. Here, stream and pollutant are stochastic variates and the problem is solved when the probability distributions of each quality parameter at each control point is given. The present model can handle several pollutants simultaneously provided that these do not interact.

Approach

For this study, a stochastic process formulating the mechanisms of dilution, dispersion, and self-purification is required. The river is divided into a finite number of reaches which have similar hydrologic and physical characteristics. These reaches can be thought of as a series of reservoirs, the outflow of one being the inflow of the next one downstream. Each reach can be visualized as a "queue" where the water must be held temporarily before continuing to the next reach, and the new concepts of queueing theory, discussed in Appendix B of this chapter, may be applied. Another tool, used in conjunction with queueing theory, is the finite Markov chain, discussed in Appendix C of this chapter.

Two different mathematical approaches may be used for the solution of complex stochastic problems, namely, discrete variable formulation and continuous variable formulation. The former has the particular advantage in this problem of being useful for the derivation of theorems and algorithms pertaining to the transport and dispersion of materials in the wash load of the stream. The latter type of formulation becomes useful when these theorems and algorithms are obtained and it is needed to simulate properly the movement of pollutants in the full range of concentrations observed in actual streams. Model SQ-SX employs the former of the two approaches and was constructed in the expectation of (1) achieving a better basic understanding of the details of the processes involved, and (2) deriving some theorems and generalizations needed for the computer simulation model.

In Model SQ-SX both flow and pollutant are treated as discrete variates. In order to prevent the matrices, which formulate the transport process, from becoming excessively large, it is necessary that only a relatively few levels of flow and pollutant be used.

The relation between the outflow from a reach and the inflow and storage of that reach is formulated as a Markov process. The outflow is a given function of the sum of the storage plus current inflow.

It was decided to study the hydrology, pollutants, and oxygen concentration in that order, for the water flow affects the pollution and the pollution affects the oxygen balance and not conversely.

In subsequent work the effect of benthal deposits, photosynthesis, and heat exchange on self-purification will be studied.

Assumptions and Limitations.

The stream is divided into a finite number of reaches labeled alphabetically as shown on Fig. 1. The mean, variance, and lag-one serial correlation coefficient of flow into reach A is assumed to be known from past records of flow. The pollutant in the model may be introduced into reach A in several ways, namely, (1) stochastic with or without serial correlation, (2) cyclical, or (3) any combination of these. The concentration of pollutant is determined in each reach from a knowledge of (1) the storage of water and (2) storage of pollutant in each reach. The mathematical techniques necessary for this calculation will be demonstrated in later sections of this chapter.

This model is a mathematical description process (technological function) rather than a decision model, but will later be modified to become a decision making model and constructed to maximize a benefit function.

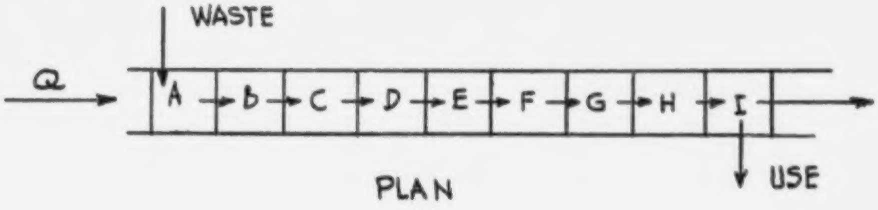


Fig. 1

(1) Given

(a) Levels of inflow of water in time, $\Delta t = 0, 1, 2, 3$

(b) Levels of storage of water = 0, 1, 2

(c) Inflow hydrology to reach A (Assume lag-one serial correlation)

		t+1			
		0	1	2	3
	0	.35	.40	.25	
	1	.125	.50	.225	.15
t	2	.267	.267	.40	.066
	3		.40	.40	.20

Stationary probability distribution
.20
.40
.30
.10

Fig. 2. Transition Matrix P

(d) Inflow of pollutant to reach A (Assume random, no serial correlation)

$$\begin{array}{cc|c|c|c|c|c|}
 & & & t+1 & & \multicolumn{2}{c}{} \\
 & & 0 & 1 & 2 & \multicolumn{2}{c}{} \\
\cline{3-5}\cline{7-7}
 & 0 & 1/3 & 1/3 & 1/3 & \multicolumn{1}{c|}{} & 1/3 \\
\cline{3-5}\cline{7-7}
t & 1 & 1/3 & 1/3 & 1/3 & \multicolumn{1}{c|}{} & 1/3 \\
\cline{3-5}\cline{7-7}
 & 2 & 1/3 & 1/3 & 1/3 & \multicolumn{1}{c|}{} & 1/3 \\
\cline{3-5}\cline{7-7}
\end{array}$$

Stationary probability
distribution

Fig. 3. Transition Matrix P

(2) Assumptions

(a) Outflow of a reach during any time period is a function of the sum of storage plus current inflow. In the example it is assumed that the function is linear, i.e., $q_B = 2/3\ (q_A + S_A)$. This is arbitrary for the purpose of exposition. It actually depends on the physical and hydraulic characteristics of the reach, which may be natural or man-made and controlled.

(b) No benthal deposits are incorporated, but these will be put in at a later time.

(c) The effect of temperature has not yet been formulated. In the completed model, temperature may be added as a stochastic variate.

(d) Perfect mixing of pollutant is assumed to take place in each reach.

(e) A one-season scheme is used to demonstrate the model in the examples of this chapter. In this scheme, the population of inflows and the population of pollutant inputs are the same during each successive time period. However, the model can be used with a two-season scheme to simulate, for example, periods of high and low runoff.

(f) Evaporation and bank storage have been neglected.

(g) $\overline{V} = \frac{\Delta x}{\Delta t}$ is approximately constant over the discrete range of flows except when $q = 0$.

Listed above are the assumptions and limitations of the present model, and an indication of those limitations that will be resolved at a later stage. The model at its present stage does, however, include three types of hydrology: (1) deterministic, (2) stochastic with and without serial correlation, and (3) cyclical, and any combination of these in order to simulate floods, droughts, trends, peaks, oscillations, and so forth. The model also includes the determination of the probability distribution of (1) the total amount of storage of flow and pollutant in any reach at any time, (2) the flow and flux of pollutant in any reach at any time.

Hydrology

The flow out of any reach during the time interval, Δt, is assumed, for this example, to be equal to two-thirds the sum of the storage in the reach and the inflow to that reach during the time interval, Δt. Expressed algebraically,

$$q_{B,t+1} = 2/3\ (S_{A,t} + q_{A,t+1}) \tag{3-1}$$

and graphically,

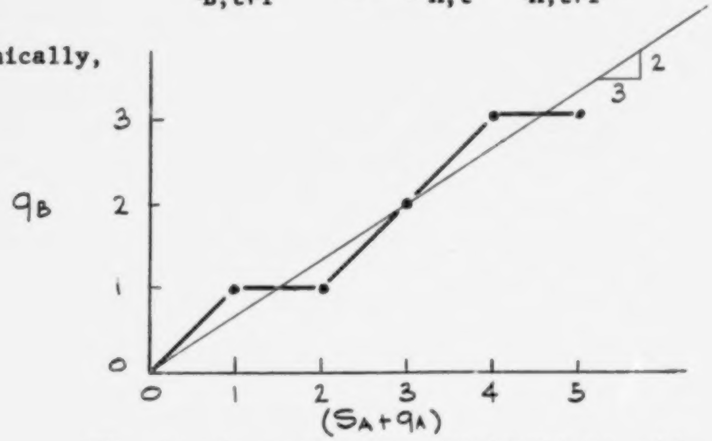


Fig. 4. Assumed Storage Yield Relationship.

For this model, the flow levels are necessarily made discrete, and therefore the ordinary rounding rules of decimal fractions must be applied. Figure 4 indicates how discrete points are utilized to approximate a continuous straight line function with slope of 2/3.

The inflow hydrology to reach A is given in Figure 2 of this chapter under Assumptions and Limitations.

The parameter along the ordinate indicates the level of flow at time t, while that across the abscissa gives the level at a time t+1. This matrix is known as the transition matrix or conditional probability matrix, for if the probability distribution of flows at time t is known, the probability distribution at time t+1 can be determined by multiplying the transition matrix P by the initial distribution vector p_i , as discussed in Appendix C, "Finite Markov Chain." Any row vector in the matrix is itself a probability distribution and the sum of the elements in any row is equal to one. If all the elements of the matrix are equal, there will be no serial correlation.

Reading across the first row of the transition matrix for hydrology, Figure 2, indicates that if the flow is zero at time t, the probability of the flow being zero at time t+1 is 0.35; the probability of the flow being one at time t+1 is 0.40; the probability of $q_{A,t+1} = 2$ is 0.25; and $p(q_{A,t+1} = 3)$ is = 0. The same applies to an initial flow level of 1, 2, and 3.

The stationary probability distribution represents the overall probability of the flow being at any stated level. Mathematically, the stationary probability vector may be thought of as the unique vector W such that WP = W, as discussed in Appendix C of this chapter. It is determined algebraically as follows: Assume a 3 by 3 transition matrix P with elements p_{ij} and a stationary probability vector p_i. For $p_{i,t}$ to equal $p_{i,t+1}$, the corresponding elements of each vector must be equal. For example:

$$\underbrace{\begin{array}{|c|c|c|}\hline p_1 & p_2 & p_3 \\ \hline \end{array}}_{\substack{\text{Probability Vector} \\ \text{at time, t}}} \quad \underbrace{\begin{array}{|c|c|c|}\hline p_{11} & p_{12} & p_{13} \\ \hline p_{21} & p_{22} & p_{23} \\ \hline p_{31} & p_{32} & p_{33} \\ \hline \end{array}}_{\text{Transition matrix}} \quad = \quad \underbrace{\begin{array}{|c|c|c|}\hline p_1 & p_2 & p_3 \\ \hline \end{array}}_{\substack{\text{Probability Vector} \\ \text{at time, t+1}}}$$

It follows from ordinary matrix algebra that:

$$p_1 = \sum_{i=1}^{3} p_i \cdot p_{i1}$$

$$p_2 = \sum_{i=1}^{3} p_i \cdot p_{i2} \qquad (3\text{-}2)$$

$$p_3 = \sum_{i=1}^{3} p_i \cdot p_{i3}$$

If the elements of the matrix p_{ij} are known, then the elements of the stationary vector $p_i = p_j$ can be determined from: (1) the set of linear equations (3-2) or (2) by an iterative technique employing the digital computer.

The matrix shown on Figure 5 is a conditional probability matrix of two interdependent stochastic parameters, namely, flow and storage. The location of the various elements in the matrix is determined by (1) the law of continuity,

$$S_{A,t} + q_{A,t+1} = S_{A,t+1} + q_{B,t+1} \qquad (3\text{-}3)$$

and (2) the assumed operation rule or storage yield relationship shown on Figure 4. Note that the description of either the ordinate or the abscissa of this transition matrix requires both the interdependent parameters, q and S. The three columns to the right of the matrix and the three rows under the matrix represent respectively (1) the sum of the storage in the reach and the inflow to the reach in the time period, Δt, $(S_A + q_A)$, determined by summing the discrete values of the two parameters on either the ordinate or the abscissa of the transition matrix, (2) the outflow from the reach, (q_B), as determined from the storage-yield relationship shown on Figure 4, and (3) the stationary probability vector, (p_S), of this transition matrix, determined by either the solution of simultaneous linear algebraic equations or applying the iterative technique as described above.

The first row of this matrix may be used as an example. The value of $(S_A + q_A)$ is equal to zero as shown in the first column to the right of the transition matrix. Figure 4 indicates that if $(S_A + q_A) = 0$, $q_B = 0$. The law of continuity states that

$$S_{A,t} + q_{A,t+1} = S_{A,t+1} + q_{B,t+1} \tag{3-3}$$

$$0 = S_{A,t+1} + 0$$

From this, it follows that the storage at the end of the time period is equal to zero. Note that the first row of the hydrology transition matrix shown on Figure 2 is located in the first row and columns 1-4 of the matrix on Figure 5, indicating that if $S_{A,t} + q_{A,t+1} = 0$, then $S_{A,t+1} = 0$.

For a second example, arbitrarily select row eleven of the transition matrix on Figure 5. As before, the first column to the right of this matrix exhibits the value of $(S_A + q_A)$ being equal to 4. If $(S_A + q_A) = 4$, $q_B = 3$ from Figure 4. From the law of continuity, $4 = S_{A,t+1} + 3$, and the storage at the end of the time period, $S_{A,t+1}$, is equal to 1. It is evident from this that the third row of the transition matrix in Figure 2 should be placed in the eleventh row and columns 5-8 of the matrix on Figure 5. Note that columns 5-8 of the matrix on Figure 5 indicate a final storage, $S_{A,t+1}$, of one.

All the rest of the elements in the transition matrix of storage and flow shown on Figure 5 are determined row by row similarly as described in examples one and two above. After this transition matrix for storage and inflow to a reach has been determined, the next step is to determine the stationary probability distribution, p_s, displayed as the third column to the right of the transition matrix on Figure 5.

This transition matrix, along with the stationary vector, is key to the entire mathematical process. The stationary probabilities of both (1) storage, and (2) outflow, can now be determined from Figure 5 as follows:

(1) Storage (S_A)

$$\begin{aligned} p\ (S_A = 0) &= a+b+c+d \\ p\ (S_A = 1) &= e+f+g+h \\ p\ (S_A = 2) &= i+j+k+l \end{aligned} \tag{3-4}$$

(2) Outflow (q_B)

$$\begin{aligned} p\ (q_B = 0) &= a \\ p\ (q_B = 1) &= b+c+e+f+i \\ p\ (q_B = 2) &= d+g+j \\ p\ (q_B = 3) &= k+h+l \end{aligned} \qquad (3\text{-}5)$$

The conditional probability matrix for outflow of reach A (which is the inflow to reach B) can also be produced by first determining the joint probability matrix shown on Figure 6. Consider the three columns to the right of the matrix on Figure 5 as containing values of the indicated parameters at time, t, and the three rows under the matrix,containing the values at time t+1. If it were desired to determine the joint probability of the outflow (q_B) being equal to zero at time t and also equal to zero at time t+1, one must "search" the second column to the right of the matrix on Figure 5 labeled $q_{B,t}$, and note all the elements that are equal to zero. Similarly, one must "search" the second row under the matrix labeled $q_{B,t+1}$, and consider all of the elements of this row that are equal to zero. Note that there is only one value of zero in either the second column or row mentioned above and that "zero" occupies the first position in the vector. The probability of $q_B = 0$ occurring at time t is "a" (i.e., $p(q_{B,t} = 0) = a$). Therefore the conditional probability of $q_{B,t+1}$ being zero if $q_{B,t}$ was zero is given as the element in the first row and first column of the transition matrix Figure 5, and is equal to 0.35 (i.e., $p_{ij}(q_{B,t+1} = 0 \mid q_{B,t} = 0)$, read the probability of $q_{B,t+1}$ being equal to zero if $q_{B,t}$ was equal to zero, $= p_{11} = .35$). To determine the joint probability of both events happening, (i.e., $q_{B,t} = 0$ and $q_{B,t+1} = 0$) one must multiply the conditional probability by the initial probability. Stated algebraically where $k = 1$ and $\ell = 1$,

$$p_{k,\ell}(q_{B,t}=0 \text{ and } q_{B,t+1}=0) = \sum_{j=1} \sum_{i=1} p_i(q_{B,t}=0) \cdot p_{ij}(q_{B,t+1}=0 \mid q_{B,t}=0) \qquad (3\text{-}6)$$

or, in other words, $p_{11}(q_{B,t} = 0 \text{ and } q_{B,t+1} = 0) = a \cdot (.35) = 0.35a$.

This is in fact the value in the first row and first column of the joint transition matrix shown on Figure 6. As a second example, one may arbitrarily decide to determine the joint probability of the occurrence of say $q_{B,t} = 2$ and $q_{B,t+1} = 1$. Again, one must "search" the second column to the right of the matrix on Figure 5 for all the values of 2. The "2's" exist in positions 4, 7, and 10 with the probabilities of d, g, and j respectively. Also, "search" the second row under the matrix for all the values of 1. The "1's" are located in the positions 2, 3, 5, 6, and 9. The necessary conditional probabilities can be found in the q_A transition matrix (Figure 5) as the elements p_{ij}, $i = 4, 7, 10$, $j = 2, 3, 5, 6, 9$ or, in other words, the elements $p_{4,2}, p_{4,3}, p_{4,5}, p_{4,6}, p_{4,9}, p_{7,2}, p_{7,3} \cdot\cdot \cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot p_{10,9}$.

As before, to determine the joint probability of the two events occurring together, such as the events $q_{B,t} = 2$ and $q_{B,t+1} = 1$, one must multiply the individual probabilities of $q_{B,t} = 2$ occurring by the corresponding conditional probabilities that if $q_{B,t}$ occurs, $q_{B,t+1}$ will occur, and then sum the results. Stated algebraically, where for this example, $k = 3$ and $\ell = 2$:

$$p_{k,\ell}(q_{B,t}=2 \text{ and } q_{B,t+1}=1) = \sum_{j=2,3,5,6,9} \sum_{i=4,7,10} p_i(q_{B,t}=2) \cdot p_{ij}(q_{B,t+1}=1 \mid q_{B,t}=2) \tag{3-7}$$

$$\text{or } p_{3,2}(q_{B,t} = 2 \text{ and } q_{B,t+1} = 1) = d(0) + d(0) + d(0) + d(.40) + d(0)$$
$$+ g(0) + g(0) + g(.267) + g(.267) + g(0) + j(0) + j(0)$$
$$+ j(.125) + j(.50) + j(0)$$
$$= + .40d + .267g + .267g + .125j + .50j$$
$$= .40d + .534g + .625j$$

This is the value to be placed in the third row and second column of the joint transition matrix shown on Figure 6.

The remaining elements of the joint matrix on Figure 6 are determined in a similar fashion as explained above. In the end, however, it is the conditional transition matrix (for inflow to reach B) that is needed and not the joint transition matrix as shown on Figure 6. This was only an intermediate step.

If the joint transition matrix of two events can be determined by multiplying the conditional probability matrix by an initial probability vector, then the reverse is true and the conditional probability matrix can be determined from the joint transition matrix by dividing it by the stationary probability vector.

Expressed algebraically:

$$\text{if} \quad p(A)\cdot p(B|A) = p(A \text{ and } B)$$

$$\text{then} \quad p(B|A) = \frac{p(A \text{ and } B)}{p(A)} \tag{3-8}$$

The conditional transition matrix shown on Figure 7 can be determined as follows: Given the joint probability matrix on Figure 6 and the stationary probability vector for inflow into reach B (q_B) as equations (3-5), divide the joint matrix by the stationary vector element by element as indicated on Figure 7 or

$$p(q_{B,t+1} \mid q_{B,t}) = \frac{p(q_{B,t+1} \text{ and } q_{B,t})}{p(q_{B,t})} \tag{3-9}$$

As an example, assume a 3x3 joint matrix M, and a stationary probability vector W.

$$M = \begin{array}{|c|c|c|}\hline m & n & o \\ \hline p & q & r \\ \hline s & t & u \\ \hline \end{array} \qquad\qquad W = \begin{array}{|c|}\hline w \\ \hline x \\ \hline y \\ \hline \end{array}$$

The conditional probability matrix $\underline{P}$ given W is determined as follows:

$$\underline{P} = \frac{M}{W} = \begin{array}{|c|c|c|}\hline \frac{m}{w} & \frac{n}{w} & \frac{o}{w} \\ \hline \frac{p}{x} & \frac{q}{x} & \frac{r}{x} \\ \hline \frac{s}{y} & \frac{t}{y} & \frac{u}{y} \\ \hline \end{array}$$

Once the conditional probability matrix for inflow to reach B, Figure 7, has been determined, this entire mathematical procedure is started again for reach B as was done for reach A just described. This may be carried on reach after reach in the same manner.

After the hydrology has been described properly in the model, the next step is to introduce a pollutant.

Pollutant

The routing of a pollutant through a river system is somewhat more complicated than hydrology because (1) the transport of the pollutant does not proceed independently but is a function of the hydrology, (2) no theoretical upper boundary can be set on the level of storage of a pollutant in a particular reach of the river as can be set for the hydrology; however, a practical limit may be fixed, (3) no upper boundary can be set on the flow (flux) of pollutant from reach to reach as could be done for the flow of water, and (4) the pollutant may decay, whereas the total amount of water was assumed to remain constant.

In this case, the matrices used in the numerical calculations must not only be made functions of storage and flow of pollutants, but must also contain a description of the hydrology. Also, these matrices will extend out in two directions appearing limitless, but must be "cut off" at some practical limit as will be shown. All the mathematical details necessary in setting up the matrices and the calculations involved are similar as for the hydrology except where noted below.

The storage-yield rule for the pollutant is given in Figure 8 for all levels of flow except when no flow occurs and at that time no pollutant will flow either. Note that this is the same linear 2/3 rule as assumed for the hydrology. This obtains from the original assumption that the pollutant is completely mixed throughout each individual reach and therefore the percentage of pollutant leaving during each time interval must be the same as that for water.

The pollutant is assumed to enter reach A, (x_A), mix with the pollutant already in storage in that reach X_A , decay at a specified rate, and then leave the reach as outflow, (x_B), according to the storage yield relationship, Figure 8.

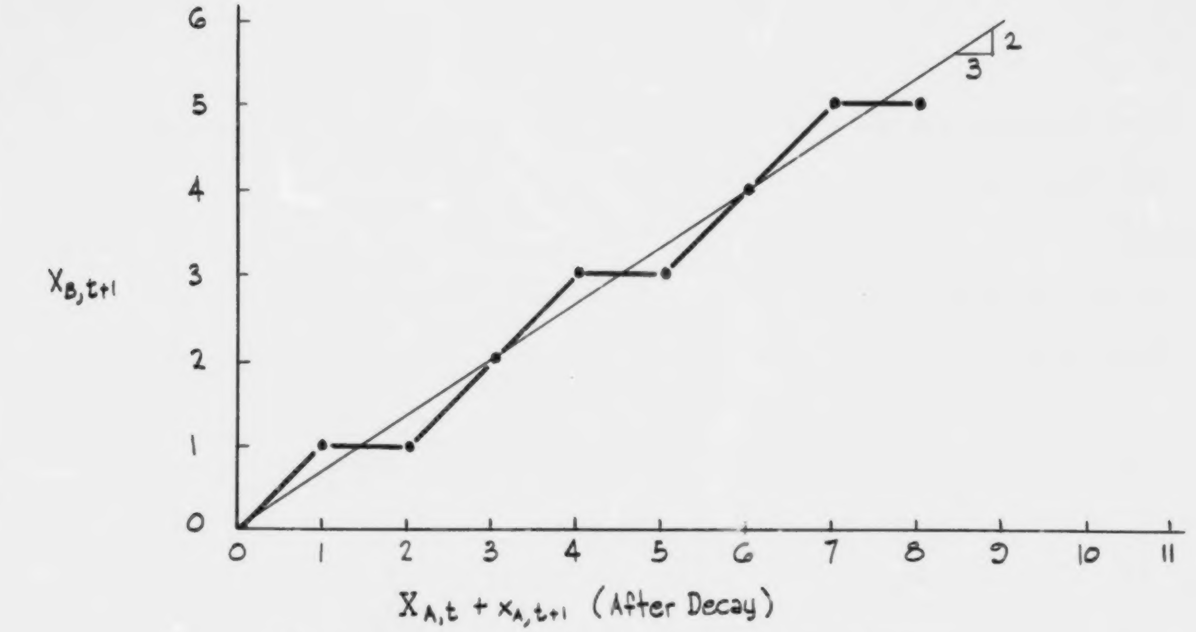


Fig. 8. Storage-Yield Relationship for Pollutant (Except at q = 0)

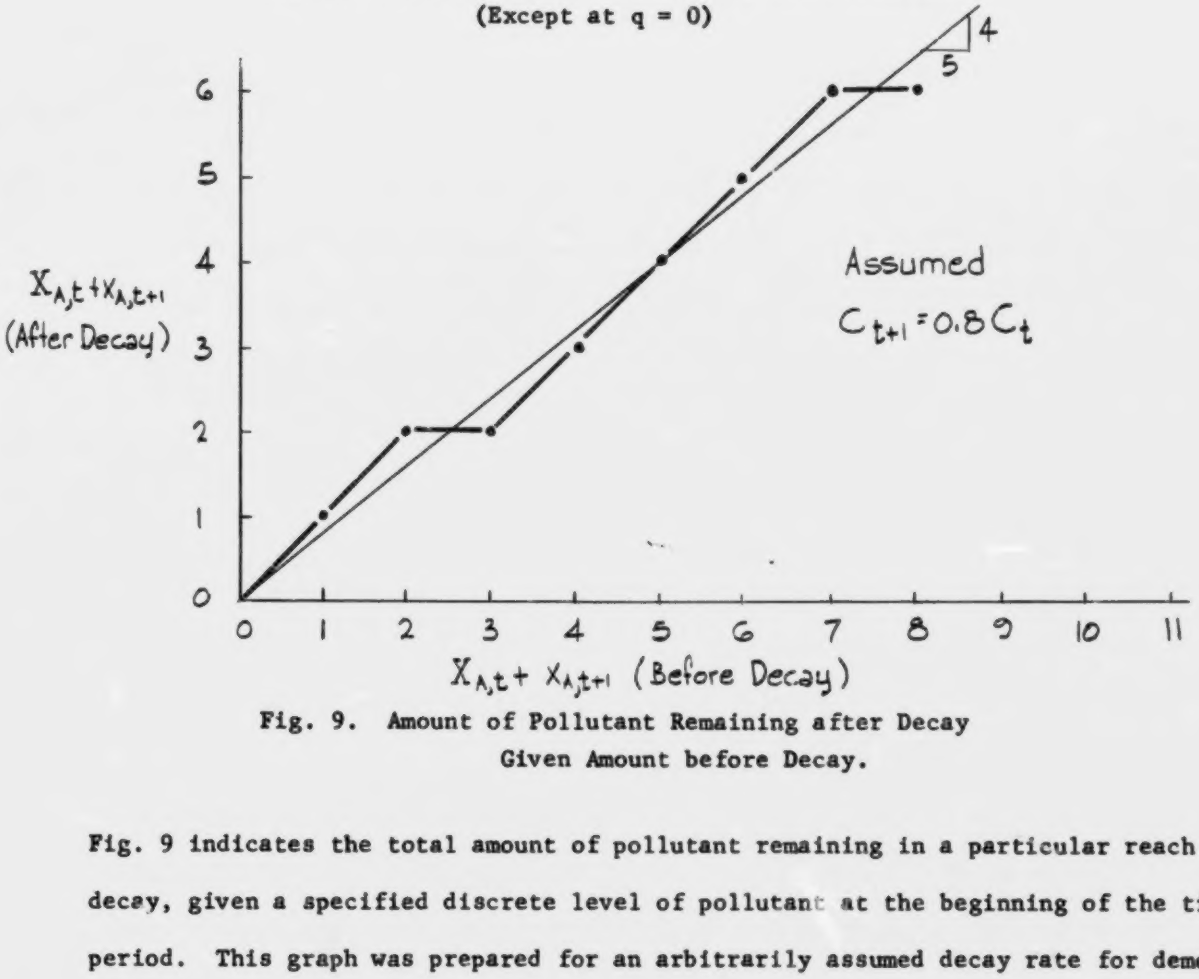


Fig. 9. Amount of Pollutant Remaining after Decay Given Amount before Decay.

Fig. 9 indicates the total amount of pollutant remaining in a particular reach after decay, given a specified discrete level of pollutant at the beginning of the time period. This graph was prepared for an arbitrarily assumed decay rate for demonstration purposes only. Here it is assumed that only 0.8 of the total pollutant remains after the specified time period, Δt.

To determine the flux of pollutant continuing to reach B, (x_B), one must first determine the amount available in reach A before decay ($X_{A,t} + x_{A,t+1}$), go to Figure 9 and determine the amount remaining after decay, proceed to Figure 8 with the value of ($X_{A,t} + x_{A,t+1}$) after decay and thus determine the available flux, X_B.

Transport of Pollutant Assuming It to Be Independent of Hydrology

If the transport of the pollutant were not dependent on the hydrology, the transition matrix for storage and flow of pollutant, given the transition matrix for inflow of pollutant to reach A on Figure 3, would be as shown on Figure 10. The first column to the right of this matrix represents the storage plus inflow of pollutant ($X_A + x_A$) before decay. The second column represents this value after decay, the third column represents the outflow of pollutant, x_B, from reach A, the fourth the storage of pollutant remaining at the end of the time period, and finally the fifth column represents the stationary probability distribution of the individual states. The only remaining difference between the techniques used to solve this problem and those used for the hydrology is the method of making this pollutant matrix finite.

The hydrology transition matrix, Figure 5, was necessarily finite whereas the pollutant matrix, Figure 10, appears to be infinite in two directions. However, for a solution, this matrix must be made finite. It is only the upper left-hand corner of a transition matrix that is of interest but just how much of it is of interest must be determined by a numerical technique. This is done by assuming a matrix may be twice the size as judged necessary. Determine the stationary probabilities of this finite matrix, and then sum the probabilities starting with the number one element and continuing in order. When the last element has been added that causes the probabilities to sum to say 0.99, stop and arbitrarily cut the matrix at this size. In other words, find the smallest value of n so that

$$\sum_{s=1}^{n} p_s = 0.99 \quad \text{where 0.99 is arbitrary} \tag{3-10}$$

Cut the existing matrix shown on Figure 10 to the proper nxn matrix. It is evident from this that the elements in the matrix beyond the nxn size are of little value, for they do not alter the elements of the stationary probability distribution beyond the n^th^ value.

After the matrix has been "cut" and the stationary probabilities determined, the remaining calculations are similar to the hydrology calculations. The stationary distribution for storage of pollutant X_A in reach A, including decay, is as follows:

$$\begin{aligned}
p\,(X_{A,t} = 0) &= a + b + c \\
p\,(X_{A,t} = 1) &= d + e + f \\
p\,(X_{A,t} = 2) &= g + h + i \\
p\,(X_{A,t} = 3) &= j + k + l \\
&\;\;\vdots \\
p\,(X_{A,t} = n) &= . + . + .
\end{aligned}$$

where "n" is the last storage value, X_A , on the "cut" matrix.

The stationary probability for the flow of pollutant from reach A, x_B, is given from Figure 10:

$$\begin{aligned}
p\,(x_B = 0) &= a \\
p\,(x_B = 1) &= b + c + d + e + f + g + h + j + \ldots \\
p\,(x_B = 2) &= i + k + \ldots \\
p\,(x_B = 3) &= l + \ldots \\
&\;\;\vdots \\
p\,(x_B = W) &= \ldots\ldots
\end{aligned}$$

where W is the largest flux of pollutant moving from reach A to reach B as indicated on the "cut" matrix of Figure 10.

The transition matrix for the inflow of pollutant to reach B, $(x_{B,t+1} | x_{B,t})$, is determined similarly as was demonstrated for the transition matrix for the inflow of water to reach B.

Transport of Pollutant with Dependence on Hydrology

The transport of a pollutant is of course in actuality dependent on the hydrology and the only reason the preceding section along with Figure 10 was added to this report is for didactic purposes. The new matrix will have to contain the hydrologic parameters, flow q and storage S, as well as the storage and flux of the pollutant shown on Figure 10. This matrix, described by the variates, X_A, x_A, q_A, and S_A, is shown on Figure 11. Note that the hydrologic transition matrix, Figure 5, as coded on Figure 12, is superimposed on and multiplied by each element in the matrix of Figure 10. The construction of the matrix on Figure 10 was based entirely on the storage-yield relationship of Figure 8. However, the new matrix on Figure 11 is based both on the storage-yield relationship of Figure 8 and the fact that the flow of pollutant, x_B, is zero when the flow of water, q_B, is zero.

The necessary calculations for the solution of this problem are similar to those discussed in the section on hydrology and the first part of this section. After the stationary distribution for storage of pollutant, X_A, the stationary distribution for the flux of pollutant, x_A, and the transition matrix $(x_{B,t+1} | x_{B,t})$ have been determined for reach A, reach B, reach C and so on are determined in a similar manner.

Note that in general the elements of the matrix on Figure 11 are located in the same relative positions as on Figure 10 except when $q_B = 0$ $(S_A + q_A = 0)$ and those elements are shifted to a new location consistent with the continuity relationship

$$X_{A,t} + x_{A,t+1} = X_{A,t+1} + x_{B,t+1} \qquad \text{or}$$

$$X_{A,t+1} = (X_{A,t} + x_{A,t+1}) \text{ after decay} \qquad (3\text{-}11)$$

Summary and Future Work

Model SQ-SX, at the present stage, includes three types of hydrology: (1) deterministic, (2) stochastic with and without serial correlation, and (3) cyclical and any combination of these in order to simulate floods, droughts, oscillations, and so forth. The model also yields the probability distribution of (1) storage of water in any reach and (2) the storage of pollutant in any reach. From these the probability distribution of the various levels of concentration of pollutants can be determined. In addition, the model incorporates a decay mechanism to handle both conservative and non-conservative wastes.

The routing of the hydrology has been programmed in Fortran automatic coding language for the computer. The program contains a main routine of approximately 70 Fortran statements and four sub-routines with a total of about 500 statements. The type of input hydrology to reach A, the maximum level of flow in the river, the maximum level of storage in each reach, and the slope of the storage-yield relationship on Figure 4 can all be varied at will without alteration of the existing program.

The Fortran program for the routing of the pollutant is still in the preliminary stages and because of this the results and conclusions for the build-up and concentration of pollutant in the various reaches are limited.

Work will continue on this model and will include:

(1) Development of a more realistic storage-yield relationship for hydrology (i.e., $q_B = f(S_A, q_A, S_B)$ rather than the existing $q_B = f(S_A, q_A)$).

(2) Construct for benthal deposits.

(3) Construct for oxygen balance.

(4) Construct for heat balance (thermal pollutant) and the effects of temperature on the various elements of the system.

(5) Effects of photosynthesis on oxygen balance.

(6) Development of a multi-season model so that stream problems with more than two seasons may be studied.

After the mechanisms and constructs proposed above are incorporated into the present mathematical model and the entire process has been programmed for solution on the computer, the model will be ready to provide the valuable information necessary for generalizations and theorems.

The next step in the evolution of this research will be the development of a mathematical construct to describe a river basin system for which data are available. Computer simulation will be used to investigate the stochastic processes of interest in this case.

APPENDIX A

Standard Notation

$S_{A,t}$	Storage of water in Reach A at time t
$q_{A,t+1}$	Inflow of water to Reach A during the time interval t to t+1
$q_{B,t+1}$	Outflow of water from Reach A during the time interval t to t+1
$X_{A,t}$	Storage of pollutant in Reach A at time t
$x_{A,t+1}$	Inflow of pollutant to Reach A during the time interval t to t+1
$x_{B,t+1}$	Outflow of pollutant from Reach A during the time interval t to t+1
a_{ij}	An element from i^{th} row and j^{th} column of any matrix A
P	Transition matrix of probabilities
W	Stationary probability vector
p_{ij}	An element from the i^{th} row and j^{th} column of the transition matrix P
p_i	A probability vector
$p(q_A = 0)$	Read as follows: the probability of q_A being equal to zero
$p(q_{A,t+1} = 0 \mid q_{A,t} = 3)$	Read as follows: the probability of q_A being equal to zero at time t+1 given that $q_A = 3$ at time t

APPENDIX B

General Discussion of Queueing Theory[1]

A "queue" is a waiting line and is commonly observed to form before ticket offices, cafeterias, bus stops, and reservoirs. Therefore, to have a queue, something must arrive at a service facility where it is often caused to wait.

Queueing theory is a branch of applied mathematics employing notions and concepts from the field of stochastic processes. It has been developed in an attempt to predict random demands from observational data and to enable a business establishment to provide adequate service for its customers with tolerable or controlled waiting. An example of controlled waiting or a controlled queue is the "service" provided by a reservoir to dampen the stochastic elements of the inflow hydrology so that the drastic effects of floods and droughts can be minimized or at least reduced appreciably.

One of the important problems arising from overpopulation is the ability of people to live with congestion. Just how one should organize congestion situations to minimize the resulting losses in time and inefficient operation or, in the case of water resource systems, how one should regulate flow (output) with the aid of a controlled queue (reservoir) to minimize losses incurred by floods and droughts is a problem in queueing theory.

The theory of queues is a descriptive theory and is not directly concerned with optimization. Rather, it attempts to simulate various queueing situations so that one may explore, understand, predict, and compare these varied conditions and thus indirectly optimize the objective function. It is used for the purpose of better understanding of queues and for the sake of introducing remedies.

Some general types of queueing phenomena to which queueing theory has applicability are in the fields of (1) communications traffic (telephone, telegraph, post office); (2) transportation traffic (air, land, sea); (3) queueing for service (theaters, restaurants, buses, hospitals, and clinics); (4) inventories and industrial processes (maintenance, assembly lines, machine interference); (5) epidemic processes in biology; (6) population growth; (7) psychological flow of nervous impulses; and (8) physical processes (operations of a set of dams, particles moving through a hole). Probably the most significant queue of all is the life queue, i.e., birth-life-death cycle with life as the waiting aspect.

This paper, however, is concerned only with the physical processes and, in particular, the simultaneous queueing of water and pollutants as they pass reach by reach through a river system. In this particular case, the output of one queue (reach A) is the input of the next queue (reach B); therefore a system is specified which consists of some finite number of queues all in series.

The finite Markov chain is a very useful mathematical tool for describing and simulating queueing problems. See Appendix C, "Finite Markov Chain."

---

1. Saaty, Thomas L., Elements of Queueing Theory, McGraw-Hill Book Co., Inc., New York, 1961.

APPENDIX C

Finite Markov Chains[2]

Basic Concepts

A Markov chain is a process that moves successively through a set of states $s_1, s_2, \ldots s_r$. In this discussion, it will be assumed that $r \geq 2$. Given that the process is in state s_i , it moves on to the next stop at state s_j with the probability p_{ij}. These probabilities can be exhibited in the form of a transition matrix P:

$$P = \text{Time}=t \quad \left[\begin{array}{c|ccccccc}
 & \multicolumn{7}{c}{\text{Time}=t+1} \\
 & S_1 & S_2 & S_3 & \cdot & \cdot & \cdot & S_r \\
S_1 & P_{11} & P_{12} & P_{13} & \cdot & \cdot & \cdot & P_{1r} \\
S_2 & P_{21} & P_{22} & P_{23} & \cdot & \cdot & \cdot & P_{2r} \\
S_3 & P_{31} & P_{32} & P_{33} & \cdot & \cdot & \cdot & P_{3r} \\
\cdot & \cdot & \cdot & \cdot & & & & \cdot \\
\cdot & \cdot & \cdot & \cdot & & & & \cdot \\
\cdot & \cdot & \cdot & \cdot & & & & \cdot \\
S_r & P_{r1} & P_{r2} & P_{r3} & \cdot & \cdot & \cdot & P_{rr}
\end{array}\right] \qquad \text{(C-1)}$$

The entries of P are nonnegative and the sum of the components in any given row is equal to 1. A vector with non-negative components having the sum 1 is called a probability vector. Therefore, each row of P is a probability vector.

To determine a Markov chain completely, one must specify the transition matrix along with a vector describing the starting state. The state vector at any time (t+1) can be produced by multiplying the transition matrix P by the state vector at time (t).

$$S_{j(t+1)} = S_{i(t)} \cdot P_{ij} \quad \text{[*]} \qquad \text{(C-2)}$$

The matrix multiplication can be repeated in this manner time unit after time unit until the state vector at time (t-1) equals the state vector at time (t). When this condition exists, the state vector is called the stationary state vector because it is no longer a function of time. A stationary state vector can be determined for all regular Markov chains.

---

[*] The operation of matrix multiplication and other related operations are described in standard textbooks such as reference number 2.

A Markov chain is called a regular chain if and only if some power of the original transition matrix contains all positive elements.

Theorem 1. If P is a transition matrix for a regular chain, then:

(1) The powers P^n approach a matrix T (that is, each entry of P^n approaches the corresponding entry of T).

(2) Each row of T is the same probability vector W.

(3) The components of W are all positive.

Theorem 2. If P is a transition matrix for a regular chain, and T and W are as in Theorem 1, then the vector W is the unique probability vector such that WP = W.

A Markov chain is called an ergodic chain if it is possible to go from every state to every other state. A regular chain is necessarily ergodic but an ergodic chain may or may not be regular. For example, if from a given state it is possible to go only to certain states in an even number of steps and to other states in an odd number of steps, then all powers of the transition matrix will contain 0's.

Theorem 3. If P is a transition matrix for an ergodic chain, then:

(1) There is a unique probability vector W such that WP = W.

(2) All components of W are positive.

---

[2] Kemeny, J. G., Mirkil, H., Snell, J. L., Thompson, G. L., Finite Mathematical Structures, Prentice-Hall, Inc., N.J., 1959, Chaps. 4 and 6.

TIME t+1

TIME t	$S_{A,t+1}$	0				1				2				at t		
$S_{A,t}$	$q_{A,t+1}$ \ $q_{A,t+2}$	0	1	2	3	0	1	2	3	0	1	2	3	S_A+q_A	q_B	p_S
0	0	.35	.40	.25										0	0	a
	1	.125	.50	.225	.15									1	1	b
	2					.267	.267	.40	.066					2	1	c
	3						.40	.40	.20					3	2	d
1	0	.35	.40	.25										1	1	e
	1					.125	.50	.225	.15					2	1	f
	2					.267	.267	.40	.066					3	2	g
	3						.40	.40	.20					4	3	h
2	0					.35	.40	.25						2	1	i
	1					.125	.50	.225	.15					3	2	j
	2					.267	.267	.40	.066					4	3	k
	3										.40	.40	.20	5	3	ℓ
at t+1	S_A+q_A	0	1	2	3	1	2	3	4	2	3	4	5			
	q_B	0	1	1	2	1	1	2	3	1	2	3	3			
	p_S	a	b	c	d	e	f	g	h	i	j	k	ℓ			

FIG. 5 CONDITIONAL PROBABILITY MATRIX OF FLOW AND STORAGE IN REACH A

	$q_{B,t+1}$				
$q_{B,t}$	0	1	2	3	P_s
0	.35a	.40a + .25a			a
1	.125b + .35e	.50b+.225b +.267c + 267c +.40e +.25e +.125f + .50f +.35i +.40i	.15b + .40c + .225f + .25i	.066c + .15f	b+c+e +f+ i
2		.40d + .267g +.267g + .125j +.50j	.40d + .40g +.225j	.20d + .066g +.15j	d+g+j
3		.40h +.267k +.267k	.40h +.40k +.40ℓ	.20h +.066k +.60ℓ	h+k+ℓ

FIG. 6 JOINT PROBABILITY MATRIX OF FLOW INTO REACH B

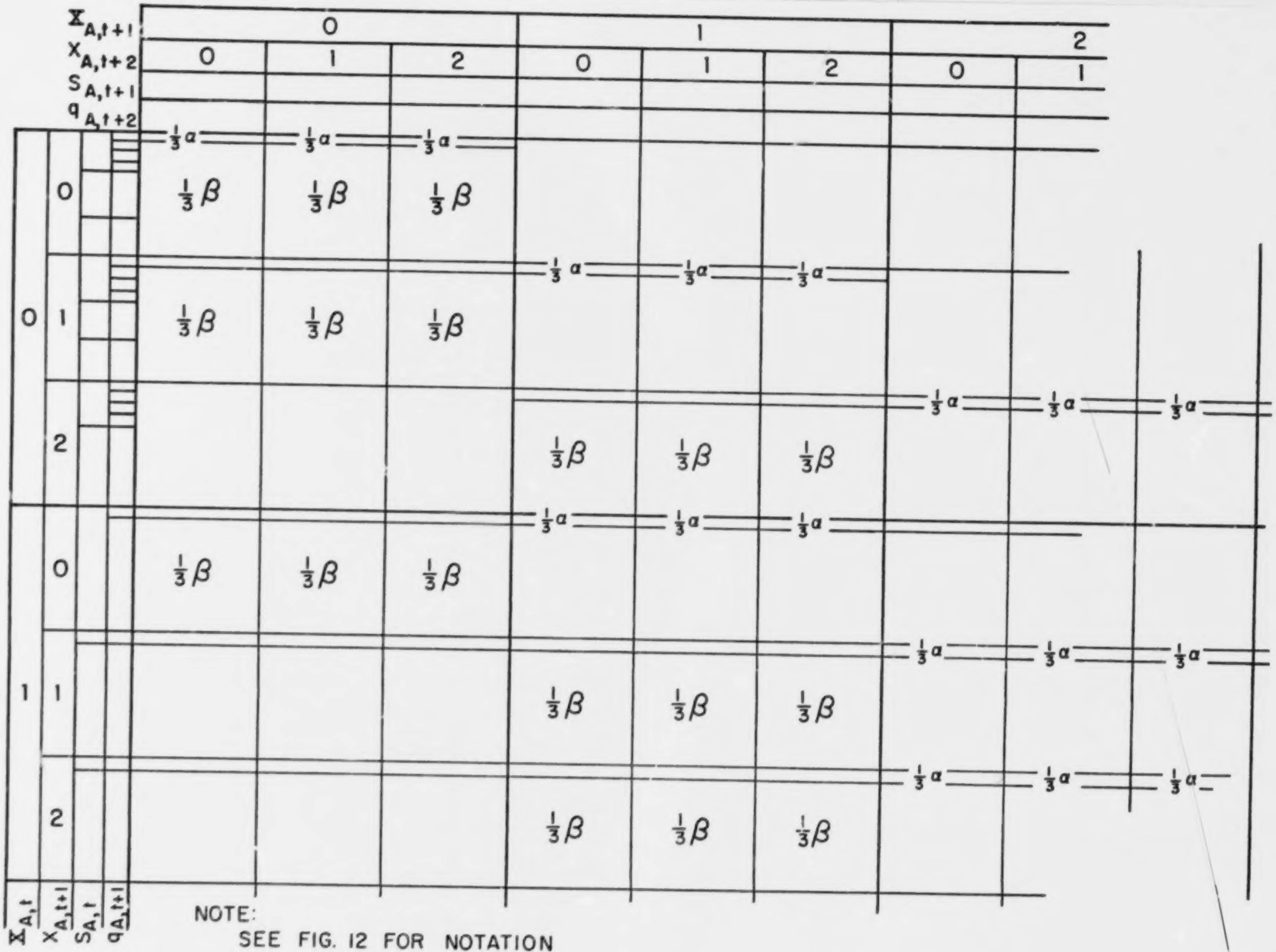
$$q_{B,t+1}$$

$$\begin{array}{c|c|c|c|c|c}
q_{B,t} & 0 & 1 & 2 & 3 & \Sigma \\
\hline
0 & .35\,\frac{a}{a} & .65\,\frac{a}{a} & & & 1.0 \\
\hline
1 & \frac{.125b+.35e}{b+c+e+f+i} & \text{etc.} & \text{etc.} & \frac{.066c+.15f}{b+c+e+f+i} & 1.0 \\
\hline
2 & & \text{etc.} & \frac{.40d+.40g+.225j}{d+g+j} & \frac{.20d+.066g+.15j}{d+g+j} & 1.0 \\
\hline
3 & & \frac{.40h+.267k+.267k}{h+k+\ell} & \frac{.40h+.40k+.40\ell}{h+k+\ell} & \frac{.20h+.066k+.60\ell}{h+k+\ell} & 1.0
\end{array}$$

FIG. 7 TRANSITION MATRIX OF FLOW INTO REACH B

		TIME t+1																			
		0			1			2			3			4			AT t			t+1	
$\bar{X}_{A,t}$	$X_{A,t+1}$ →	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	$(\bar{X}_A + X_A)$	DECAY	X_B	$\bar{X}_A$	p_s
TIME t 0	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$													0	0	0	0	a
	1	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$													1	1	1	0	b
	2				$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$										2	2	1	1	c
1	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$													1	1	1	0	d
	1				$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$										2	2	1	1	e
	2				$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$										3	2	1	1	f
2	0				$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$										2	2	1	1	g
	1				$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$										3	2	1	1	h
	2				$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$										4	3	2	1	i
3	0				$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$										3	2	1	1	j
	1				$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$										4	3	2	1	k
	2				$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$										5	4	3	1	ℓ

FIG. 10 TRANSITION MATRIX FOR POLLUTANT ASSUMING POLLUTANT INDEPENDENT OF HYDROLOGY

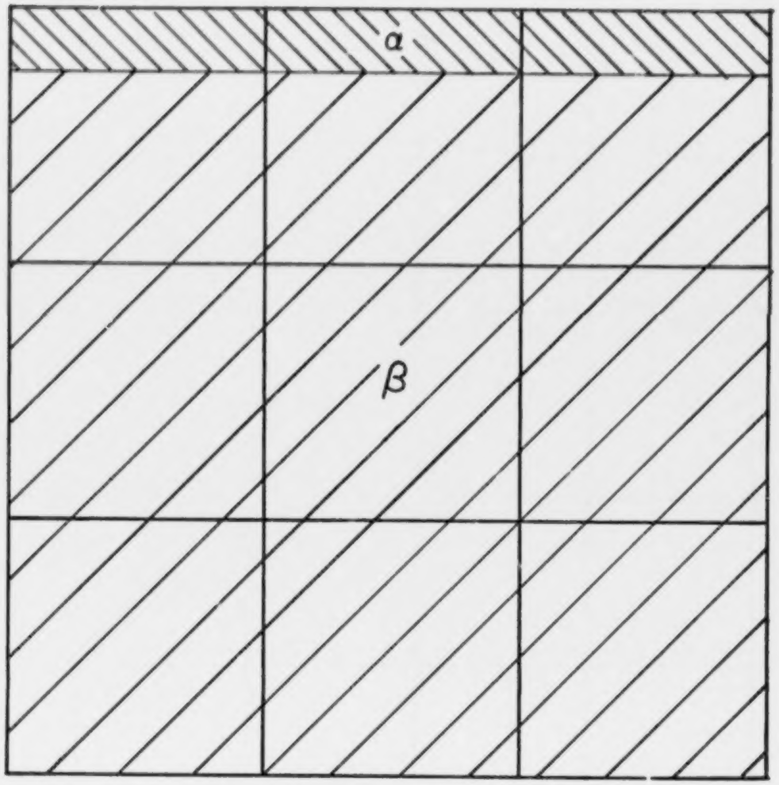

$\mathbb{X}_{A,t+1}$
$X_{A,t+2}$
$S_{A,t+1}$
$q_{A,t+2}$
0
1
2
0 1 2 0 1 2 0 1
$\frac{1}{3}\alpha$
$\frac{1}{3}\beta$
$\mathbb{X}_{A,t}$
$X_{A,t+1}$
$S_{A,t}$
$q_{A,t+1}$

NOTE:
SEE FIG. 12 FOR NOTATION

FIG. 11 TRANSITION MATRIX FOR POLLUTANT WITH DEPENDENCE ON HYDROLOGY

			t + 1											
		$S_{A,t+1}$	0				1				2			
		$q_{A,t+2}$	0	1	2	3	0	1	2	3	0	1	2	3
	0	0	.35	.40	.25									
		1	.125	.50	.225	.15								
		2					.267	.267	.40	.066				
		3						.40	.40	.20				
	1	0	.35	.40	.25									
t		1					.125	.50	.225	.15				
		2					.267	.267	.40	.066				
		3						.40	.40	.20				
	2	0					.35	.40	.25					
		1					.125	.50	.225	.15				
		2					.267	.267	.40	.066				
		3										.40	.40	.20
	$S_{A,t}$	$q_{A,t+1}$												

=



MATRIX SQ.

NOTE:
α CASE WHERE $q_{B=0}$
β CASE'S WHERE $q_{B>0}$

FIG. 12 SIMPLIFIED NOTATION FOR CONDITIONAL PROBABILITY MATRIX OF FLOW AND STORAGE SHOWN ON FIG. 5

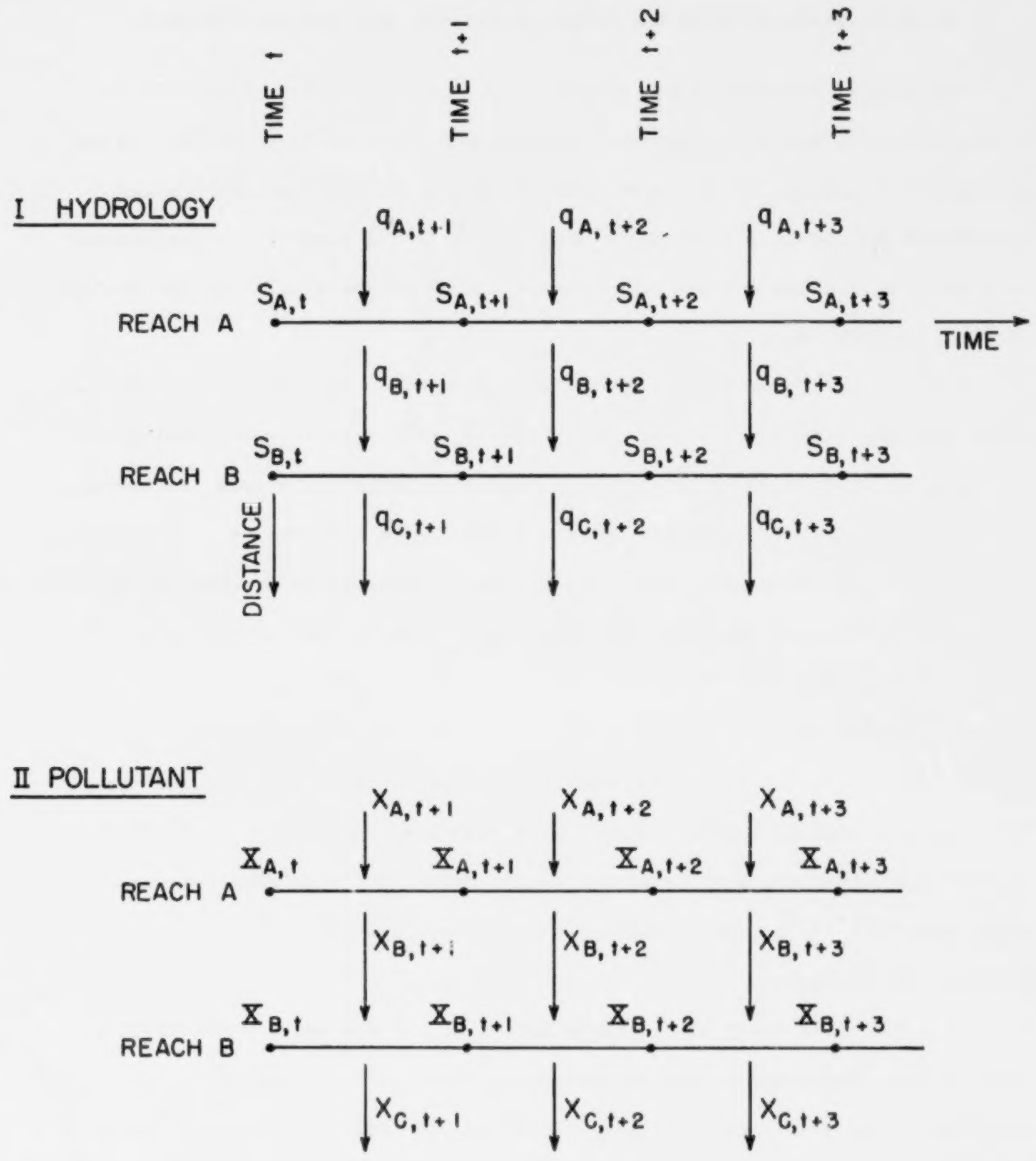

TIME t
TIME t+1
TIME t+2
TIME t+3
I HYDROLOGY
$q_{A,t+1}$
$q_{A,t+2}$
$q_{A,t+3}$
$S_{A,t}$
$S_{A,t+1}$
$S_{A,t+2}$
$S_{A,t+3}$
REACH A
TIME
$q_{B,t+1}$
$q_{B,t+2}$
$q_{B,t+3}$
$S_{B,t}$
$S_{B,t+1}$
$S_{B,t+2}$
$S_{B,t+3}$
REACH B
DISTANCE
$q_{C,t+1}$
$q_{C,t+2}$
$q_{C,t+3}$
II POLLUTANT
$X_{A,t+1}$
$X_{A,t+2}$
$X_{A,t+3}$
$\mathbf{X}_{A,t}$
$\mathbf{X}_{A,t+1}$
$\mathbf{X}_{A,t+2}$
$\mathbf{X}_{A,t+3}$
REACH A
$X_{B,t+1}$
$X_{B,t+2}$
$X_{B,t+3}$
$\mathbf{X}_{B,t}$
$\mathbf{X}_{B,t+1}$
$\mathbf{X}_{B,t+2}$
$\mathbf{X}_{B,t+3}$
REACH B
$X_{C,t+1}$
$X_{C,t+2}$
$X_{C,t+3}$

FIG. A-1 STANDARD NOTATION

CHAPTER FOUR

An Application of Decision Theory to Sanitary Engineering Problems

Statistical decision theory presents an interrelated set of methods for making decisions embodying physical and economic uncertainties, and errors and inadequacies of data; it is the science of ranking alternative designs and operations and choosing a course of action. It is the purpose of this chapter to show in an exploratory way how decision theory can be applied to the problem of water quality management.

A water quality management system of a stream may be taken to include the river together with possibilities for affecting or controlling the quality of its water at various use points by dams, water purification plants, and sewage and industrial wastes treatment plants. Depending on the physical configuration of the river and use points, some optimal mix of investments in these three possibilities obtains for specified objectives. The optimal mix is that combination of structures and operations that most closely achieves the design objective. Modern decision theory provides a method for determining the best design. In a water quality management system the costs of sampling and analysis accruing to a quality control network or monitoring program must be balanced against the potential benefits of systematic use of the resulting data. Furthermore, decisions as to how to employ the data are functions of what kinds and how many are available.

At a different level of the optimization of the system, a waste treatment plant design, for example, may be considered in detail. In the so-called complete treatment plant with separate sludge digestion, the optimal investment pattern among the primary sedimentation subsystem, the secondary subsystem (biological unit, such as trickling filter, followed by settling tank), and the sludge digestion subsystem is not necessarily obvious. In fact the complicated interactions between these subsystems, together with economic considerations, have forced conventional practice toward unit-by-unit designs that seem quite independent of each other. Performance of sedimentation basins, however, does affect the performance of other units in the plant; and design rules suggesting detention times, for example, do reflect an estimate of what is optimal for the plant as a whole.

Errors due to lack of precision in the measurement of design parameters, such as population or per capita production of sewage, give rise to uncertainty that leads to unoptimal designs and operating policies. The traditional approach is to invoke factors of safety in ways that reflect rational concern for uncertainty and the cost of reducing it. In decision theory the increase in costs, or more generally, the loss in benefits attributable to uncertainty is frequently called regret. In problems concerning cost minimization, regret is the increase in cost of a design based on erroneous information over the cost of the design that would have been made had the information been perfect, that is, if the design parameters had not been in error but represented the situation, or nature, exactly.

With an estimate of the distribution of possible states of nature the expected, or average, regret associated with a selected design can be computed. Expected regret is called risk. The more complete the information and the smaller the error on which the design parameters are based, the lower will be the expected regret, or risk. It may be possible by additional investment in experimentation, or collection of data, to reduce the errors in the selection of the design parameters. Concomitantly some reduction in the expected regret obtains. Said differently, with more nearly perfect information a design with lower risk can be selected. Indication of how much investment in experimentation can be justified by a decrease in expected regret and how to use the data to achieve an optimal design are the inseparable concerns of decision theory.

In introducing some of the concepts of statistical decision theory to aid in the solution of problems of water quality management, it is desirable for several reasons to choose a problem that has the following characteristics: it should be a problem of nationwide importance in which difficult administrative questions as well as technical ones arise; it should be a problem for which there exists a technological function - a known relationship between the physical parameters; it should involve measurements and observations which are subject to a large amount of error and uncertainty; and finally, it should be a problem for which there exists a large and homogenous body of up-to-date physical and economic data.

On the basis of the foregoing criteria the problem selected for the initial adaptation of decision theory is the disposal of sewage in metropolitan fringe areas. With the continuing movement of populations into the peripheral zones of cities, the problem of disposal of liquid wastes from individual households and small communities has become complex and difficult. According to the 1960 census there are 18,000,000 septic tanks in the continental United States.

Little doubt exists that the diffuse "wheel-spoke" growth of our cities will continue and that the administrative and technical problems involved in fringe area sewage disposal will increase until a rational methodology is evolved for the design of efficient and safe waste management systems. Fortunately there exists a massive collection of data pertaining to this problem from an extensive and intensive investigation at the Robert A. Taft Sanitary Engineering Center that started in 1946 and was carried forward in the laboratory, in pilot plants and in the field. Results have been reported by Weibel, Coulter, Bendixen and others.(1, 2, 3, 4, 5, 6)

Thus the model is the application of decision theory to the optimal design of systems of sewage disposal units for individual homes. The use of soil for the absorption of effluent from such pretreatment devices as septic tanks has a long tradition in sanitary engineering practice; but widely accepted rules of design were not established until the pioneer work of Ryon(7) was promulgated by the New York State Department of Health and the Joint Committee on Rural Sanitation. He was apparently the first to suggest a correlation between fresh-water percolation rates and the behavior in time of a subsoil that is absorbing a septic-tank effluent. His recommendations largely influenced the subsequent formulation of design rules and standards, temporarily culminating in the publication of the Manual of Septic Tank Practice by the U. S. Public Health Service.(8) The work of Ludwig,(9) and also Kiker,(10) should be mentioned in this connection. Recently, Thomas et al.,(11) presented a mathematical model of a soil absorption system that included pertinent technical and economic parameters derived from the definitive FHA-sponsored studies of on-lot disposal made by the U. S. Public Health Service at the Sanitary Engineering Center that have been mentioned above. With slight modification it will be presented in succeeding paragraphs for it figures importantly in the overall decision models. However, let us first state succinctly the decision problem.

Suppose we have some information about the soils in a county that are used for absorption from seepage pits or beds receiving septic-tank effluent. Also, in particular, there are data on the soils of several subdivisions within the county.* These data include absorption capacities of the soils from some of these subdivisions. This property is a bundle of measurable parameters,(11) that reflects the capacity of a soil for absorbing septic-tank effluent in terms of person-years per square foot; it is obtained from existing soil absorption systems. The data also include some fresh water percolation rates.** In some of the subdivisions

---

[*] The dichotomy, county-subdivision, is used here only for the sake of convenience in presentation. It could be a population-subpopulation classification of general applicability.

[**] At present the fresh water percolation test seems most useful; but the model described herein may be used to rate the utility of new methods yet to be perfected such as perhaps air diffusion instead of fresh water percolation. The model is generally applicable in this respect.

we may have both kinds of data, in others one kind and not the other, and in some we may have neither kind. The number of each kind of data will in general be different within a subdivision, and also from subdivision to subdivision. We now go into a new subdivision for which a certain number of dwelling units is being designed. It is the purpose of this decision model to answer the following questions. For this number of installations how do we use the existing data from other subdivisions? How many tests of both kinds should be made in the new subdivision? And what strategy or design rule uses these new observations optimally? The objective is the determination of the design that will minimize the economic risk in the new subdivision. The optimal numbers of each kind of test and the optimal design rule for using the old and the new data are not at all intuitively obvious. The decision problem is complicated because the costs of the two kinds of tests are presumably different, and the amount of information (with respect to reducing the economic risk) that each conveys is almost certainly different. And furthermore, the optimal investment in experimentation will vary with the number of dwelling units. With the general problem in mind we turn now to a mathematical model of a soil absorption system.

Technology and Economics of Soil Absorption Systems

Thomas et al.,(11) proposed that the capacity of a soil to absorb pretreated septic-tank effluent, or the effluent of some other device, could be measured in person-years per square foot. They designated this by the symbol, L .

$$L = \frac{fPt}{A} \qquad (1)$$

where P is the number of persons (expressed as cubic feet of effluent produced in one year for each person)

f is a factor that reflects the type of pretreatment (septic-tank effluent might behave differently from the effluent of an aerobic device). It may also be used to account for the effects of food waste grinders, clothes washers, etc.

fP might be considered as "equivalent persons"

t is the time in years before the soil becomes ultimately clogged

A is the nominal area being loaded (that is, the area of the trench or seepage pit considered as a rectilinear or cylindrical body).

Let the capital cost of a soil absorption field, C, be represented by a linear equation.

$$C = a + bA \tag{2}$$

or, from equation (1), write

$$C = a + \frac{bfPt}{L} \tag{3}$$

The present value of this soil absorption system, C_F, with a renewal period of t years, an interest rate of i, and an economic time horizon of T years is

$$C_F = \left[a + \frac{bfPt}{L}\right]\left\{1 + (1+\beta)\left[\frac{1 - \exp[-k(T-t)]}{\exp(kt) - 1}\right]\right\} \tag{4}$$

where $\exp(x) = e^x$; $e \doteq 2.7183$, the base of Napierian logarithms

$k = \ln(1+i)$; the force of interest, or instantaneous discount rate

β is a factor which enables one to handle the possibility that replacements cost more (or less) than the original installation. If $\beta = 0.0$, the replacement cost is equal to the first cost; if $\beta = 0.2$, it is 20% higher.

If a, b, fP, L, i, β, and T are known we could calculate that renewal period, t', which would minimize the present value of the total cost. This is conveniently done by determining the value of t when $\frac{dC_F}{dt} = 0$. The method that is used is a Newton-Raphson process which states that the $(j+1)^{th}$ approximation to t' is related to the j^{th} approximation by the following recursive formula

$$t'_{j+1} = t'_j - \frac{\left.\frac{dC_F}{dt}\right]_j}{\left.\frac{d^2C_F}{dt^2}\right]_j} \tag{5}$$

When the numerical difference between the $(j+1)^{th}$ and j^{th} approximation is acceptably small we may consider t'_{j+1} as the ideal renewal time for the given parameters. The associated minimum cost can then be computed by substituting t' and the given parameters in equation (4). For typical values of a, b, fP, β, i, and T the relationship between the ideal, or minimum, cost and soil capacities is as shown in Figure 1. The associated ideal renewal periods are shown in Fig. 2.

Suppose the true capacity was 400 cubic feet per square foot and we designed as though it were 200. It follows from equation (1) that the actual renewal period will be twice the design renewal period.

Let η be the true capacity

L_d be the design, or guessed, capacity

t' be the ideal renewal period, that is, the one we would select if we knew what the true capacity was

t'' be the design renewal period, the decision variable

t''' be the actual renewal period that would obtain if we acted as though nature was L_d and it really was η

It follows that

$$t''' = t'' \frac{\eta}{L_d} \qquad (6)$$

The present value of the cost of a system designed on the basis of L_d when we should have selected η is:

$$C' = \left[a + \frac{bfPt''}{L_d} \right] \left\{ 1 + (1+\beta) \left[\frac{1 - \exp\,[-k(T-t''')]}{\exp(kt''') - 1} \right] \right\} \qquad (7)$$

Call C' the "actual" cost. The relationship of C' and L_d is shown in Fig. 3 for different values of η. In the parlance of statistical decision theory the difference between the "actual" cost associated with a given design, L_d, when nature, or the true value, is η, and the ideal cost if $L_d = \eta$ is called regret.

It is an index of the "regret" we would experience if we failed to estimate the state of nature correctly, that is, we underdesigned or overdesigned with a concomitant increase in the cost over what it would have been had we guessed the true capacity. The regret for various design values and true values is shown in Fig. 4. The notion of risk arises when we consider the average, or expected, regret, weighted according to the expected distribution of the true capacities. Mathematically the risk associated with a given design is:

$$\text{Risk } (L_d) = \int_{-\infty}^{\infty} r\ (L_d,\eta)\ f\ (\eta)\ d\eta \tag{8}$$

where $r\ (L_d,\eta)$ is the regret associated with a design value of L_d when the true capacity is η

η may be a transform, such as a logarithm of the true value

$f(\eta)$ is the frequency distribution of η or its transform.

If there is not to be a single design but a design rule, or strategy, based on observations yet to be made, mathematical expression for risk is a double integral equation similar to equation (8) but also integrated over the expected distribution of L_d. This case will not be considered in detail in this chapter but the methodology will be outlined later.

The One-Dimensional Statistical Decision Model for Soil Absorption Systems Using Only Existing Data

Suppose there is one type of test available that can be used to estimate the true capacity, η. This is a measure of the person-years per square foot that existing systems have shown and can be gotten by such surveys as reported in the study of Seepage Beds.(5) Let this be called the y-statistic. Suppose there exists a county that is broken into k subdivisions (Fig. 5). In some of these subdivisions we have y data; in the j^{th} subdivision we have n_{yj} data that measure the y-statistic. The i^{th} measurement of this statistic in the j^{th}

subdivision can be written

$$y_{ij} = \bar{\eta} + \sigma_\eta t_j + \frac{\sigma_y}{\sqrt{n_{yj}}} u_{ij}$$

and the average of the tests in the j^{th} subdivision is

$$\bar{y}_j = \bar{\eta} + \sigma_\eta t_j + \frac{\sigma_y}{\sqrt{n_{yj}}} u_j \qquad (9)$$

where $\bar{\eta}$ is the population mean of true capacity

σ_η is the standard deviation of the subdivision true capacities (η_j) about the population mean

$\bar{\eta} + \sigma_\eta t_j = \eta_j$; that is the subdivision soil capacity

σ_y is the standard deviation of the y-statistic in the j^{th} subdivision. It may be noted here that the estimation of this parameter involves an assumption of homoscedasticity (constant variance) that may be tested by a Bartlett's statistic(12)

t and u are standard normal variates with means of zero and variances of one.

What we want is infomation about the population parameters $\bar{\eta}$ and σ_η; and we use equation (9) to secure estimates of these parameters.

The following well-known unbiased estimates are needed:

$$E\ (y_{ij}) = \bar{y}_j = \frac{\sum_{i=1}^{n_{yj}} y_{ij}}{n_{yj}} \qquad (10)$$

$$\hat{\eta}_j = \bar{y}_j \qquad (11)$$

$$E\ (\bar{y}_j) = \frac{\sum_{j=1}^{k} \bar{y}_j}{k} \qquad (12)$$

$$\hat{\sigma}_{\bar{y}}^{2} = \frac{\sum_{j=1}^{k} \left(\bar{y}_j - E(\bar{y}_j)\right)^{2}}{k - 1} \tag{13}$$

$$E(\eta_j) \equiv \bar{\eta} \tag{14}$$

$$\hat{\bar{\eta}} = E(\bar{y}_j) \tag{15}$$

$$\hat{\sigma}_{yj}^{2} = \frac{\sum_{i=1}^{n_{yj}} (y_{ij} - \bar{y}_j)^{2}}{n_{yj} - 1} \tag{16}$$

$$\hat{\sigma}_{y}^{2} = \frac{\sum_{j=1}^{k} (n_{yj}-1)\, \sigma_{yj}^{2}}{\sum_{j=1}^{k} n_{yj} - k} \tag{17}$$

Note that

$$\sigma_{\bar{y}}^{2} = E\,[(\bar{y}_j)^{2}] - [E\,(\bar{y}_j)]^{2} \tag{18}$$

From equation (15)

$$[E\,(\bar{y}_j)]^{2} = \hat{\bar{\eta}}^{2} \tag{19}$$

Square both sides of equation (9)

$$\bar{y}_j^{2} = \bar{\eta}^{2} + \sigma_{\eta}^{2}\, t_j^{2} + \frac{\sigma_y^{2}}{n_{yj}}\, u_j^{2} + 2\bar{\eta}\, \sigma_{\eta}\, t_j + 2\bar{\eta}\, \frac{\sigma_y}{\sqrt{n_{yj}}}\, u_j + 2\sigma_{\eta} \frac{\sigma_y}{\sqrt{n_{yj}}}\, u_j\, t_j \tag{20}$$

Taking the expected value of both sides yields

$$E[(\bar{y}_j)^{2}] = \bar{\eta}^{2} + \sigma_{\eta}^{2} + \frac{\sigma_y^{2}}{E(n_{yj})} \tag{21}$$

where

$$E(n_{yj}) = \frac{\sum_{j=1}^{k} n_{yj}}{k}$$

Substituting equations (19) and (21) into (18) gives

$$\sigma_{\bar{y}}^{2} = \sigma_{\eta}^{2} + \frac{\sigma_{y}^{2}}{E(n_{yj})}$$

An unbiased estimate of the variance of the η_j about the population mean, $\bar{\eta}$, is

$$\hat{\sigma}_{\eta}^{2} = \hat{\sigma}_{\bar{y}}^{2} - \frac{\hat{\sigma}_{y}^{2}}{E(n_{yj})} \tag{22}$$

We now have unbiased estimates of the first and second moments of the true distribution of soil capacities throughout the county.

A plot of observed capacities on log-probability paper leads to the assumption that a log-normal distribution will adequately represent nature. A typical log-normal plot of observations on seepage pits made by the U. S. Public Health Service in California$^{(13)}$ is shown in Fig. 6. Let the standard normal variate be defined as

$$t = \frac{\ln \eta - M}{S} \tag{23}$$

where M and S are the population mean and standard deviation of the log-normal distribution. They are related to the unbiased estimates of $\bar{\eta}$ and σ_{η} as follows

$$M = \ln\left[\frac{\hat{\bar{\eta}}}{\left(1 + \frac{\hat{\sigma}_{\eta}^{2}}{\hat{\bar{\eta}}^{2}}\right)^{1/2}}\right] \tag{24}$$

$$S = \ln\left[1 + \frac{\hat{\sigma}_{\eta}^{2}}{\hat{\bar{\eta}}^{2}}\right] \tag{25}$$

Now define a design variate, s_d

$$s_d = \frac{L_d - \hat{\bar{\eta}}}{\hat{\sigma}_\eta} \tag{26}$$

For a given $\hat{\bar{\eta}}$ and $\hat{\sigma}_\eta$ let the regret associated with a design using s_d when nature is t, be designated as $r(r,s_d)$. The equation for risk, or average regret weighted according to the distribution of states of nature, $f(\eta)$, is

$$\text{Risk } (s_d) = \int_{-\infty}^{\infty} r(t,s_d)\, f(t)\, dt$$

or

$$\text{Risk } (s_d) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} r(t,s_d) \exp\left(-\frac{t^2}{2}\right) dt \tag{27}$$

Using a multinomial approximation to the normal distribution over a range of 3 standard deviations either side of the mean we calculate the risk due to taking the action s_d when given $\hat{\bar{\eta}}$ and $\hat{\sigma}_\eta$. The results for typical values of the pertinent economic and technical parameters and an $\hat{\bar{\eta}}$ and a $\hat{\sigma}_\eta$ of 350 and 175 cubic feet per square foot, respectively, are given in Fig. 7.

A program for determining that value of s_d that minimizes the risk has been written for use on the IBM 7090 (or other Fortran compatible) computer. It is essentially a Newton-Raphson process. The results of a large number of runs with considerable permutation of the parameters over a wide range indicates the relatively strong parameters to be β and the coefficient of variation, c_v (that is, $\hat{\sigma}_\eta / \hat{\bar{\eta}}$). The results are shown in Fig. 8 which indicates the optimal s_d for different values of β for coefficients of variation in the range 0.0 to 2.0.

For a given $\hat{\bar{\eta}}$ or $\hat{\sigma}_\eta$ and β read the optimal s_d from Fig. 8. Substitute back into equation (26) to determine the optimal L_d. Take this value (and its associated optimal renewal period, t", from Fig. 2), together with the design loading, $f_d P_d$, and use equation (1) to determine the optimal design area. For given values of c_v and β the value of s_d read from Fig. 8 is valid provided that the value of bfP/aL lies within the range 0.05 to 1.5. This range is broad enough to include most design situations. Moreover, because of the broadness of this range it is evident that large errors in the cost parameters a and b and in the estimate population load fP do not appreciably affect the present decision as to the unit soil capacity (L_d) for design. Appropriate values of β will ordinarily fall within the range of 0 to 0.4 and for many regions a value of β of 0.2 will be satisfactory. Therefore, it may be concluded under the assumptions of this analysis that the most important parameter affecting s_d is c_v which reflects the heterogeneity of the soil. In regions with heterogeneous soils $(c_v > 0.5)$ it is economic to use soil design capacities that are considerably below the true mean capacity for the region. In homogeneous soils as indicated in Fig. 8 the optimal design capacities approach the mean.

Example: If the estimates of the parameters $\bar{\eta}$ and σ_η that describe the distribution of subdivision soil capacities are 300 and 180 cubic feet per square foot, respectively, as calculated from equations (14) and (22). And furthermore, $\beta = 0.2$ then from Fig. 8 the optimal $s_d = -0.357$; for c_v $\frac{180}{300} = 0.6$. The optimal $L_d = 300 + (-.357) \times 175 \triangleq 238$. From Fig. 2 the optimal design renewal period, t", for a capacity of 238 is approximately 13.9 years. If the design person-loading, $f_d P_d$, were $4 \times 2500 = 10,000$ cf/yr,[*] the optimal design area, A_d, is $10,000 \times 13.9/238 = 583$ ft^2. We repeat that this is the design that minimizes the economic risks given an estimate of the distribution of the true subdivision soil capacities. The risk per installation for the conditions given in the example together with the typical economic parameters noted in Fig. 1 is $12.71. It is a "one-dimensional" model since only one type of past field data, y-statistics, are available. It is a "no test" solution since in the new subdivision for which the installation is being designed no past data are available and no additional tests are considered.

---

[*] The average per capita production of sewage in unsewered areas is about 50 gpd.

Summary and Conclusions

A methodology for ranking alternative designs under conditions of uncertainty is being developed. The objective is to minimize economic risk. In this chapter a decision model was presented that showed how to use existing field data on soil characteristics in a county to arrive at an optimal design in a subdivision for which soil tests are not available. The solutions that were indicated in Fig. 8 are suitable for those situations where some *a priori* estimate of the true distribution of subdivision soil capacities is available but no further testing is under consideration.

If an observation (or more than one) is made, an *a posteriori* estimate of $f(\eta)$ is obtained and the decision problem is to find what design rule, $f(L_d|\bar{y})$, that is, L_d is specified for each value of $\bar{y}$, minimizes the economic risk over the expected distribution of observations. This amounts to solving the double integral equation referred to above. The next stage of research would be to balance the cost of improving the *a posteriori* distribution by further experimentation against the reduction in risk thereby gained. The amount of experimentation, or new observation, that is economically justified for a given number of dwelling units could then be specified.

A statistical correlation model is being developed to use the information from fresh water percolation rates and observations of soil capacities (person-loadings of existing systems). It will be possible with this model to specify optimal numbers of new tests of each kind and, furthermore, the optimal design rule, $f(L_d | (\bar{x} \text{ and } \bar{y}))$, that is, L_d is specified for each possible combination of $\bar{x}$ and $\bar{y}$ in such a way as to minimize the economic risk.

REFERENCES

1. S.R. Weibel, C.P. Straub and J.R. Thoman, "Studies on Household Sewage Disposal Systems, Part I," Federal Security Agency, Public Health Service, Environmental Health Center, 1949.

2. T.W. Bendixen, M. Berk, J.P. Sheehey and S.R. Weibel, "Studies on Household Sewage Disposal Systems, Part II," Federal Security Agency, Public Health Service, Environmental Health Center, 1950.

3. S.R. Weibel, T.W. Bendixen and J.B. Coulter, "Studies on Household Sewage Disposal Systems, Part III," U. S. Dept. of Health, Education, and Welfare, Public Health Service, Robert A. Taft Sanitary Engineering Center, 1955.

4. J.B. Coulter and T.W. Bendixen, "Final Report to the Federal Housing Administration on the Study to Determine if Distribution Boxes Can Be Eliminated Without Inducing Increased Failure of Disposal Fields," U. S. Dept. of Health, Education, and Welfare, Public Health Service, Robert A. Taft Sanitary Engineering Center, 1958.

5. J.B. Coulter, T.W. Bendixen and A.B. Edwards, "Report of a Study Sponsored by the Federal Housing Administration to Develop Practical Design Criteria for Seepage Beds as a Method for Disposal of Septic Tank Effluent," U. S. Dept. of Health, Education, and Welfare, Public Health Service, Robert A. Taft Sanitary Engineering Center, 1960.

6. T.W. Bendixen, A.A. McMahan, R.E. Thomas and J.B. Coulter, "Report of a Study Sponsored by the Federal Housing Administration to Determine the Effect of Food Waste Grinders on Septic Tank Systems," U. S. Dept. of Health, Education, and Welfare, Public Health Service, Robert A. Taft Sanitary Engineering Center, 1961.

7. Henry Ryon, "Notes on the Design of Sewage Disposal Works," Personal Publication, 1928.

8. "Manual of Septic Tank Practice," Public Health Service Publication 526, 1957.

9. H.F. Ludwig and J.S. Ludwig, "Equilibrium Percolation Test for Estimating Soil Leaching Capacity," Modern San., 4, 1952.

10. J.E. Kiker, Jr., "A Soil Percolation Test for Determining the Capacity of Soils for Leaching Sewage Effluents," Bulletin 23, Florida Eng. Ind. Exper. Sta., 1948.

11. H.A. Thomas, Jr., J.B. Coulter, T.W. Bendixen and A.B. Edwards, "Technology and Economics of Household Sewage Disposal Systems," J.W.P.C.F. 32, 113, 1960.

12. W.J. Dixon, F.J. Massey, Jr., Introduction to Statistical Analysis, McGraw-Hill, 1957, pp. 179-180.

13. Personal Communication.

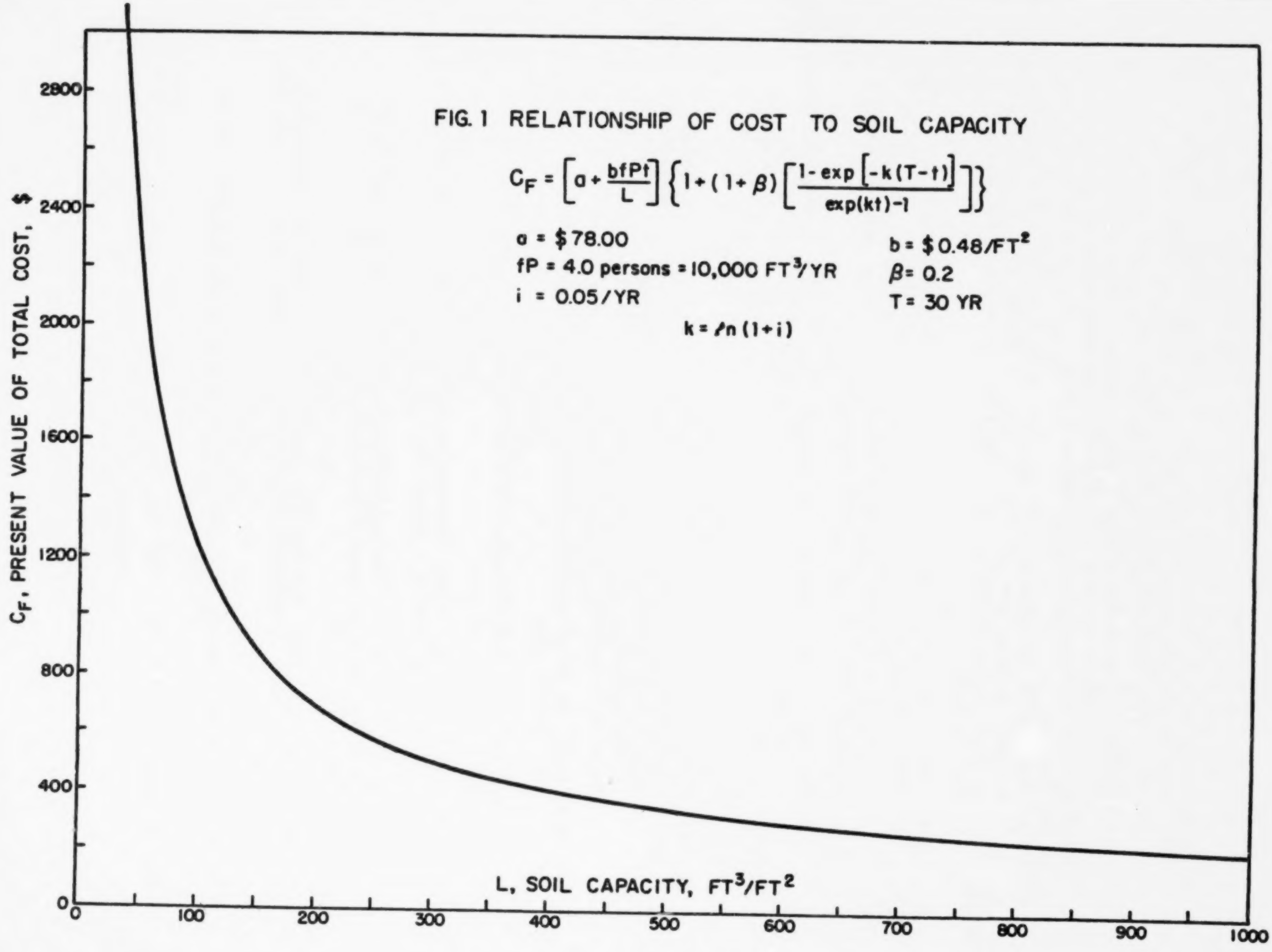


FIG. 1 RELATIONSHIP OF COST TO SOIL CAPACITY

$$C_F = \left[a + \frac{bfPt}{L}\right]\left\{1 + (1+\beta)\left[\frac{1-\exp\left[-k(T-t)\right]}{\exp(kt)-1}\right]\right\}$$

a = \$78.00
fP = 4.0 persons = 10,000 FT^3/YR
i = 0.05/YR
b = \$0.48/FT^2
β = 0.2
T = 30 YR
$k = \ln(1+i)$

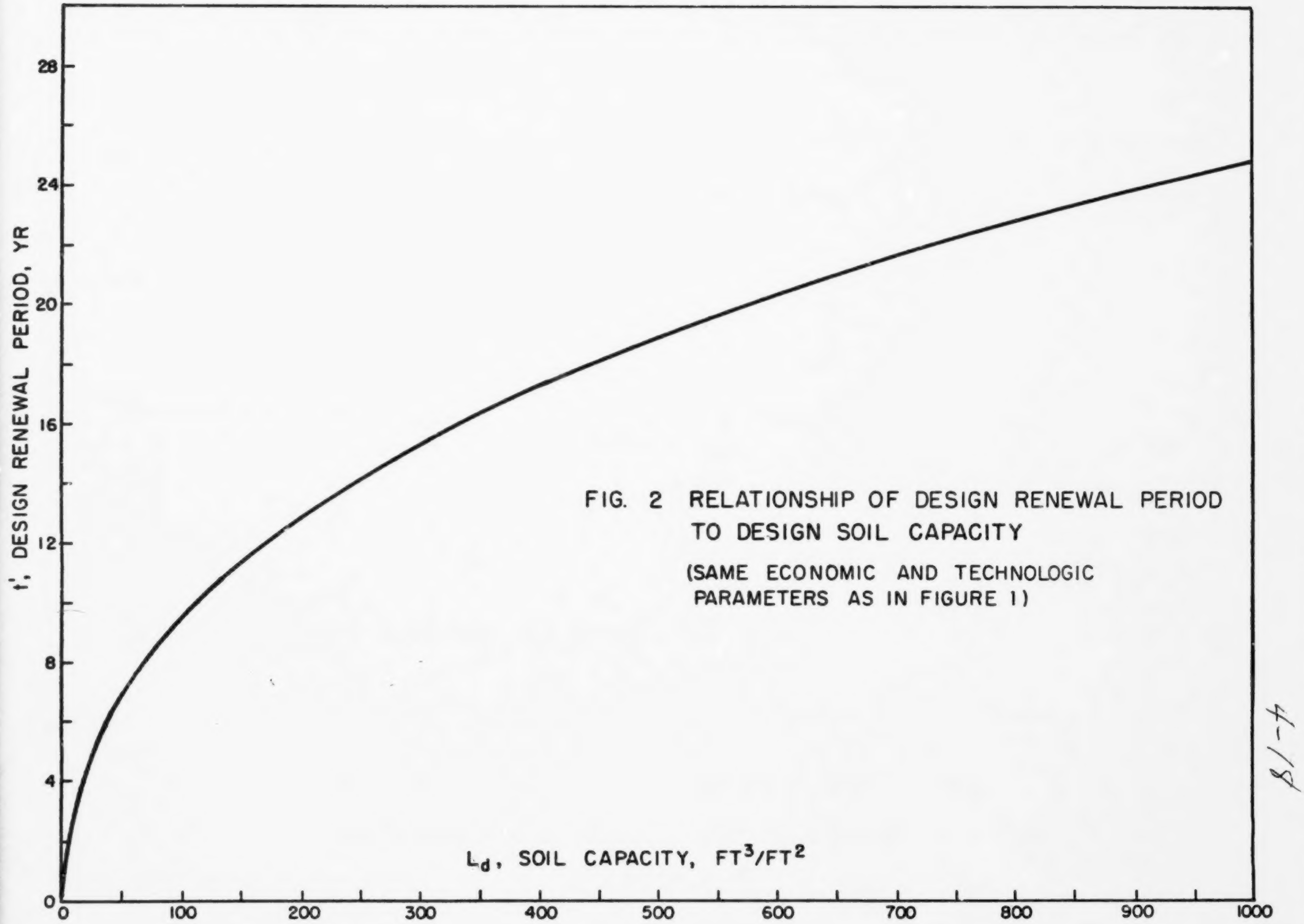


FIG. 2 RELATIONSHIP OF DESIGN RENEWAL PERIOD TO DESIGN SOIL CAPACITY

(SAME ECONOMIC AND TECHNOLOGIC PARAMETERS AS IN FIGURE 1)

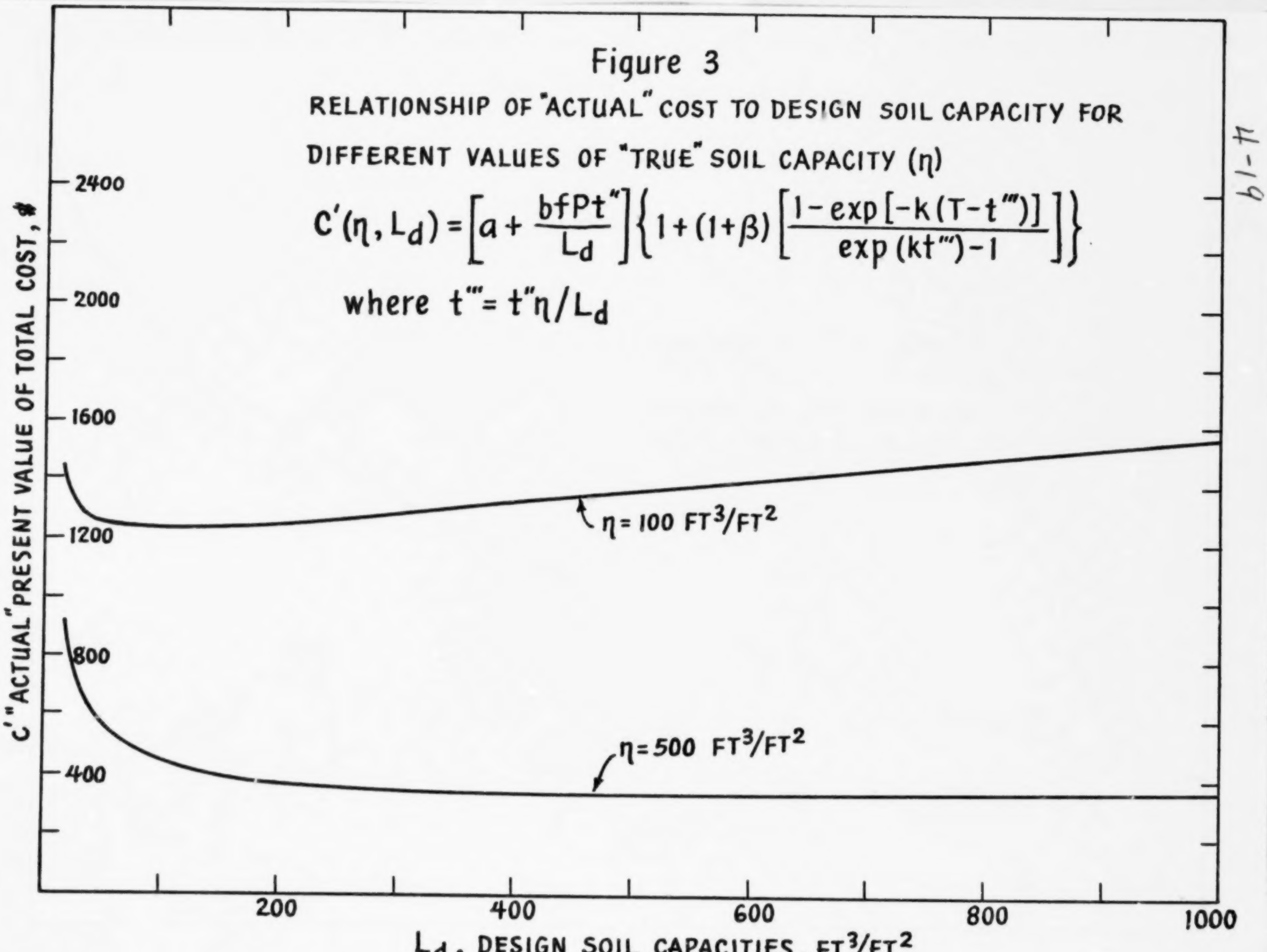


Figure 3

RELATIONSHIP OF "ACTUAL" COST TO DESIGN SOIL CAPACITY FOR DIFFERENT VALUES OF "TRUE" SOIL CAPACITY (η)

$$C'(\eta, L_d) = \left[a + \frac{bfPt''}{L_d}\right]\left\{1 + (1+\beta)\left[\frac{1-\exp[-k(T-t''')]}{\exp(kt''')-1}\right]\right\}$$

where $t''' = t''\eta/L_d$

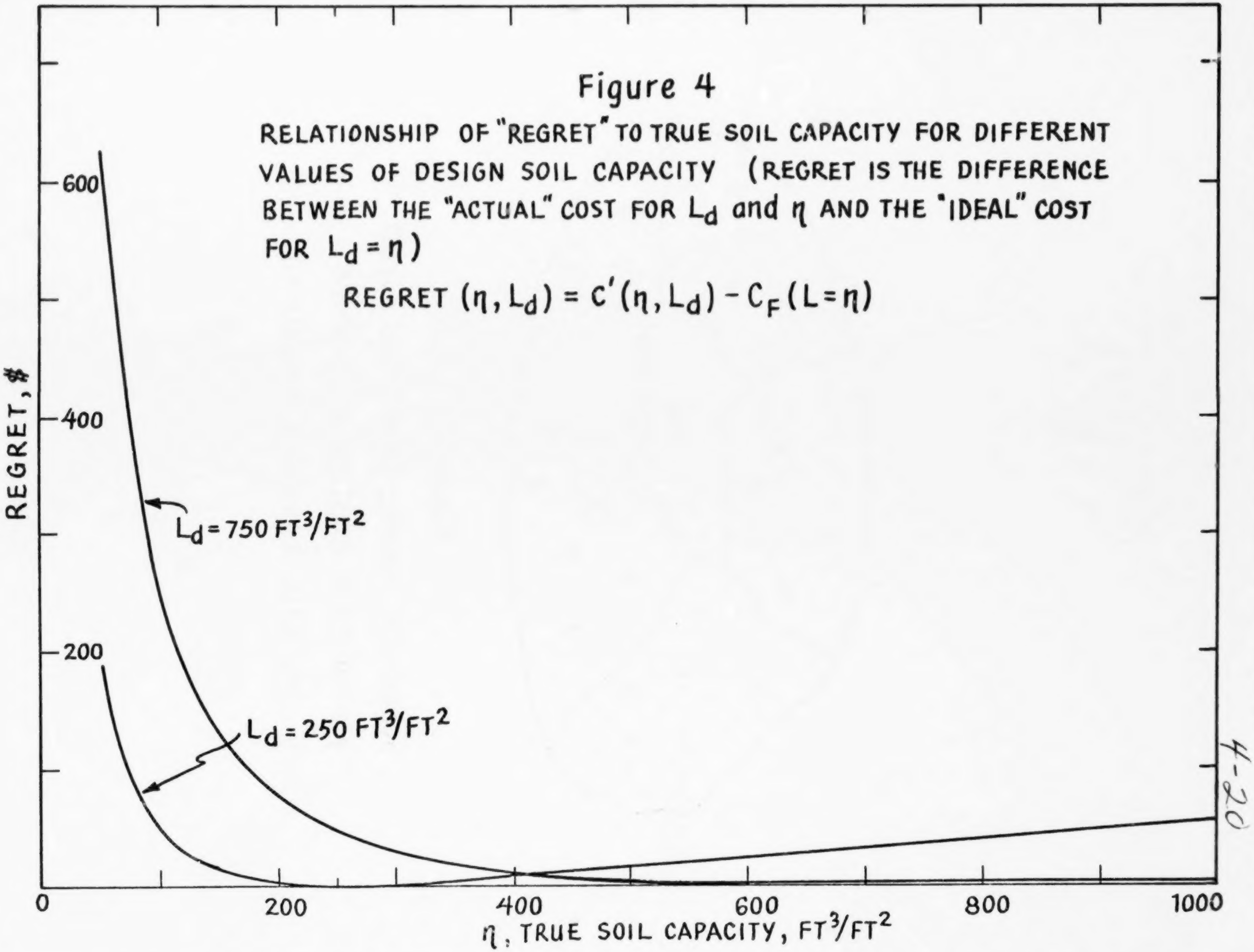


Figure 4

RELATIONSHIP OF "REGRET" TO TRUE SOIL CAPACITY FOR DIFFERENT VALUES OF DESIGN SOIL CAPACITY (REGRET IS THE DIFFERENCE BETWEEN THE "ACTUAL" COST FOR L_d and η AND THE "IDEAL" COST FOR $L_d = \eta$)

$$\text{REGRET}(\eta, L_d) = C'(\eta, L_d) - C_F(L = \eta)$$

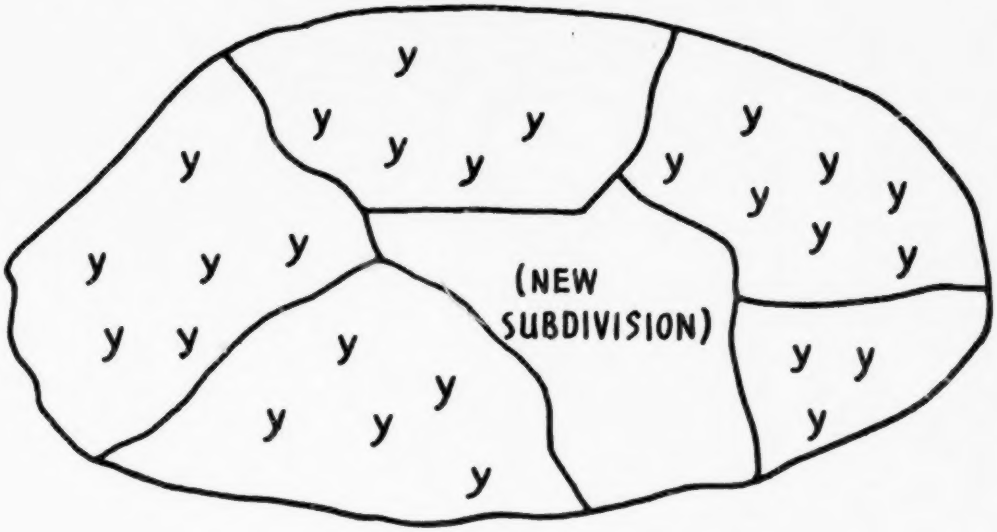

y
y
y
y
y
y
y
y
y
y
y
y
y
y
y
y
y
y
y
y
y
y
y
y
y
y
y
y
y
(NEW SUBDIVISION)

Figure 5

Schematic county-subdivision layout. y's represent existing sites where y statistic has been measured

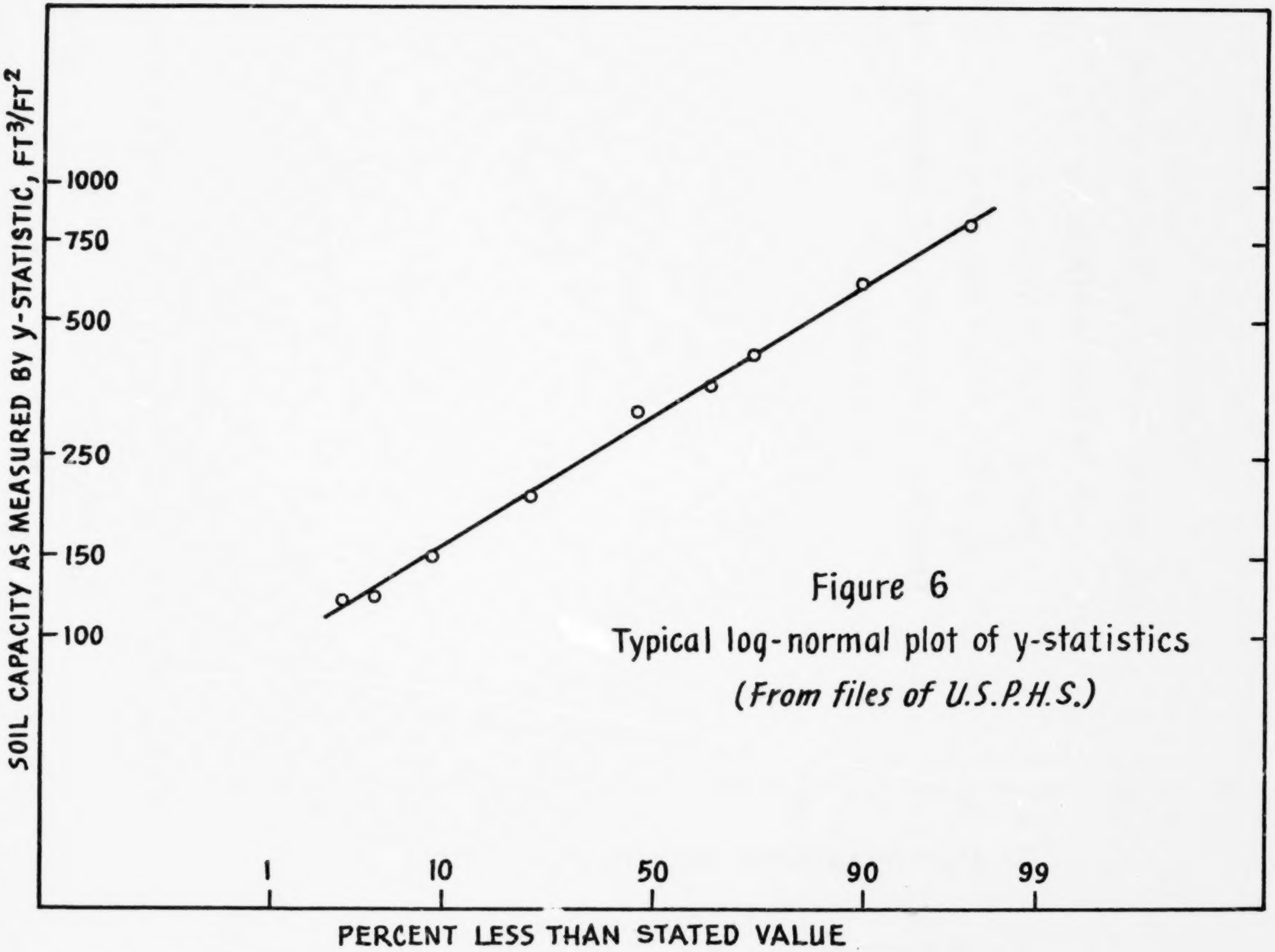


Figure 6

Typical log-normal plot of y-statistics

(From files of U.S.P.H.S.)

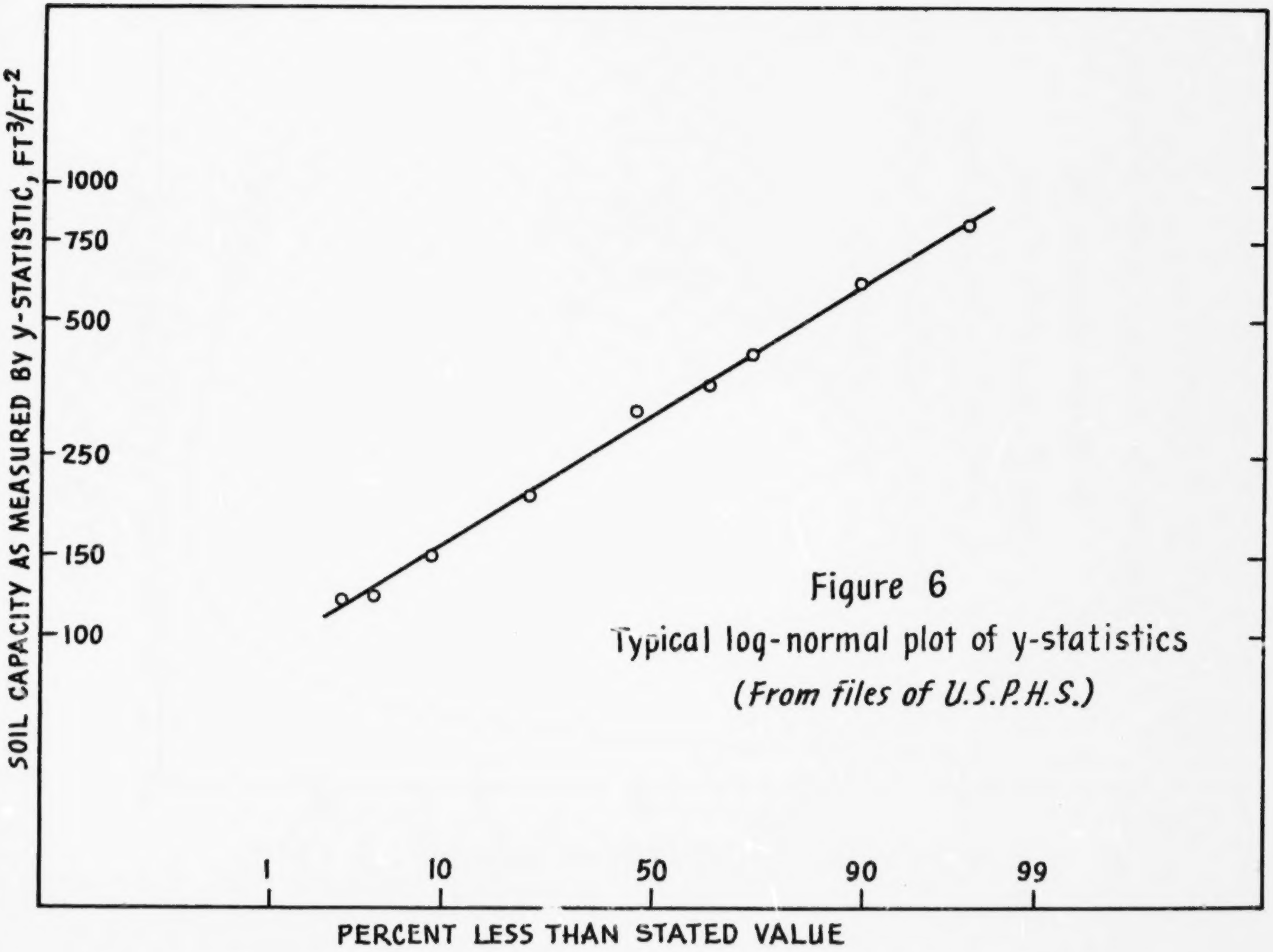


Figure 6

Typical log-normal plot of y-statistics

(From files of U.S.P.H.S.)

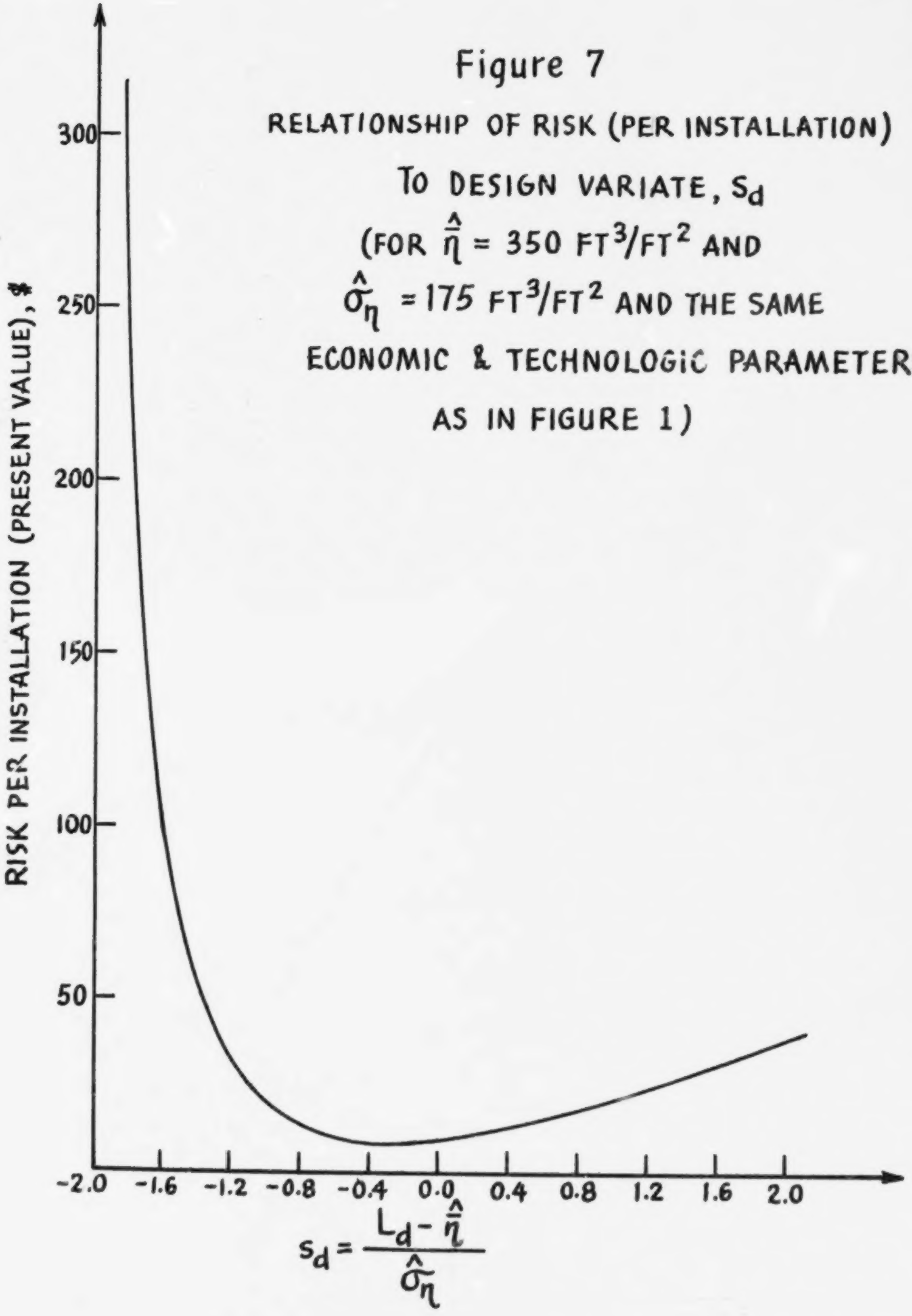
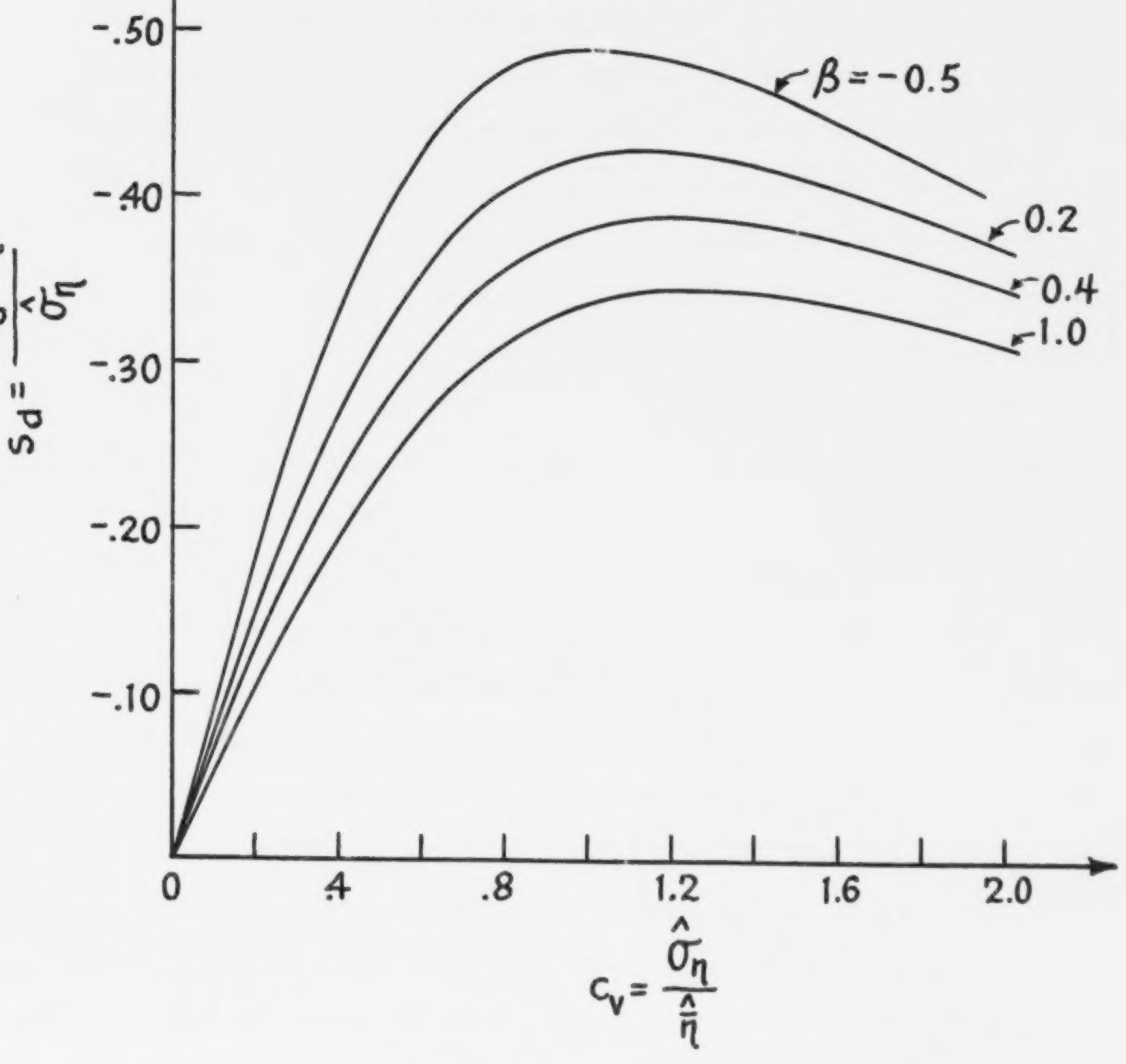


Figure 7

RELATIONSHIP OF RISK (PER INSTALLATION) TO DESIGN VARIATE, S_d (FOR $\hat{\hat{\eta}} = 350$ FT^3/FT^2 AND $\hat{\sigma}_\eta = 175$ FT^3/FT^2 AND THE SAME ECONOMIC & TECHNOLOGIC PARAMETERS AS IN FIGURE 1)

Figure 8

THE DESIGN VARIATE, s_d, VS. THE COEFFICIENT OF VARIATION, c_v FOR DIFFERENT VALUES OF β



CHAPTER FIVE

Mathematical Model for Determining the Optimal Sizes of Water Treatment Plants

Standard engineering textbooks (1)(2) list three primary considerations in establishing the design capacity of a water treatment plant:

1. The period of design
2. The population to be served during the design period
3. The rate of water use as reflected by estimates of future per capita consumption

The design capacity is then determined by multiplying the estimated future population by the estimated future water consumption. Numerous pages are devoted to presenting methods for computing both the future population and the water demand. Primary emphasis is given to physical criteria; economic considerations play a secondary role. The period of design is handled commonly with a rule of thumb that says when interest rates are high (greater than 3%), the period of design should be 10-15 years, and when it is low, 20-25 years.

Our method of determining the size of water treatment plant is an attempt to apply economic concepts on a more rigorous basis. Utility functions and benefit time-streams are used to determine the optimal design. However, in order to alleviate mathematical complexities and facilitate a reasonable introduction to the problem, several simplifying assumptions are made. The basic assumption is that the water treatment plant is the critical factor in the design. It is assumed that the raw water supply is abundant and the distribution system is adequate for both present and future needs. In addition, it is predicated that there are no budget constraints either in the amount of money available for the construction of the plant or in the time in which the investment must be repayed. This simple model presupposes a deterministic population growth*, and the utility function reflects net utility. That is, maintenance, operation, and repair of the plant have been netted-out in the utility function.

BASIC MODEL

For this basic model the population growth is assumed to be deterministic. It is further assumed that it varies arithmetically or geometrically, with time as shown in Figure 1. P_o is the population at time "zero" (i.e., now) and b is the rate of growth in terms of persons per year. Since both b and P_o are known, the population, P(t), at any time t can be found from the equation

$$P(t) = P_o + bt$$

The utility function is similar to that used by Thomas and Watermeyer (3) except that for this simple case it is assumed that the combined benefits, or utility, depends on only three parameters rather than four. Thus, if q represents the per capita-day water availability, the benefit, u, in dollars per person per day is a single-valued function of the water availability. This relation is illustrated in Figure 2, where q_B is that value of water availability at which benefits are no longer increased, u_o is the value of the benefit function at q_B, and g is utility per gallon of water. This "breakpoint" type of benefit function, although too simple perhaps for referred analyses, represents the main characteristics of utility accruing to municipal water use. It indicates that with increasing consumption, benefits increase. However, there is a maximum rate of water use (corresponding to q_B) after which no additional benefit is obtained.

---

*Currently work is being carried forward on a more elaborate model in which the population growth is treated as a stochastic variate; other modifications are being incorporated to derive a generalized formulation which will be useful in a broad range of municipal situations.

For values of q less than q_B, the utility function may be expressed by the equation

$$u_1(q) = u_o + gq - gq_B$$

For values of q greater than q_B it becomes

$$u_2(q) = u_o$$

As stated earlier, q is in terms of gallons per person per day, u in dollars per person per day, and g in dollars per gallon.

If Q_o represents the capacity of the present treatment plant (if any) in terms of gallons per day and Q_D represents the design capacity of the future plant, also in gallons per day, then the water available to each person at time t is $Q_o + Q_D$ divided by the population at time t. That is,

$$q = \frac{Q_o + Q_D}{P_o + bt}$$

If t_1 is the time at which $q = q_B$ then $q_B = \dfrac{Q_o + Q_D}{P_o + bt_1}$

so that

$$t_1 = \frac{1}{b}\left[\frac{Q_o + Q_D}{q_B} - P_o\right]$$

Since Q_o, Q_D, P_o and b are constant, as time increases q decreases. Thus for $q < q_B$, $t > t_1$ and for $q > q_B$, $t < t_1$. The utility functions thus are

$$u_1(q) = u_o + gq - gq_B \qquad t_1 < t < oo$$

$$u_2(q) = u_o \qquad o < t < t_1$$

In determining the cost function, it is assumed that the total initial cost of the plant varies linearly with the capacity of the plant. This relation is shown in Figure 3 where K is the total cost in dollars, Q_D the design capacity of the plant in terms of gallons per day, and j the cost per gallon per day capacity after a fixed initial cost, K_o.

The cost of a plant of any capacity Q_D therefore may be given by the equation

$$K(Q_D) = K_o + jQ_D$$

Thus from these considerations, it is evident that the benefit, B(t) during any year t to t + 1 is equal to the value of the utility function at year t times the population at year t.

$$B(t) = P(t) \cdot u(q)$$

The present value of an annual benefit B(t) at compound interest i, t years from now, can be found from

$$G = B(t) \quad (1 + i)^{-t}$$

making the transformation

$$e^{ht} = (1 + i)^t$$

the expression for the present worth of the benefit becomes

$$G = B(t) \; e^{-ht}$$

or

$$G = P(t) \cdot u(q) \; e^{-ht}$$

where of course $h = \ln(1 + i)$

Since the net benefit stream is the sum of the present worths of all the yearly benefits minus the initial cost of the plant, it may be represented in integral form by

$$N = \int_o^T P(t) \cdot u(q) \quad e^{-ht} \quad dt \quad -K$$

where T is the economic time horizon (i.e., that time after which the contribution of the present worth of yearly benefits to the net benefit stream is negligible. However, since we are concerned with the net benefits attributed to the new plant only, the present worth of the future yearly benefits now being provided (i.e., P_o times u_o) can be deducted. The net benefit stream due to the construction of the new plant then becomes

$$N = \int_o^T P(t)\ u(q)\ e^{-ht}\ dt - K - \int_o^T P_o u_o\ e^{-ht}\ dt$$

For Case I therefore,

$$N = \int_o^{t_1} P(t)\ u_2(q)\ e^{-ht}\ dt + \int_{t_1}^T P(t)\ u_1(q)\ e^{-ht}\ dt - K - P_o u_o \int_o^T e^{-ht}\ dt$$

where

$$t_1 = \frac{1}{b}\left(\frac{Q_o + Q_D}{q_B} - P_o\right)$$

Since $u_2(q) = u_o$ and $u_1(q) = u_o + gq - gq_B$

$$N = \int_o^t (P_o + bt)(u_o)\ e^{-ht}\ dt + \int_{t_1}^T (P_o + bt)(u_o + gq - gq_B)\ e^{-ht}\ dt \ - K$$
$$- P_o u_o \int_o^T e^{-ht}\ dt$$

But as shown earlier $q = \dfrac{Q_o + Q_D}{P_o + bt}$

so that

$$N = \int_o^{t_1} (P_o + bt) u_o e^{-ht} dt + \int_{t_1}^T [P_o(u_o - gq_B) + g(Q_o + Q_D) + b(u_o - gq_B) t e^{-ht} dt \ -K$$
$$- P_o u_o \int_o^T e^{-ht} dt$$

Integrating, recalling that

$$\int e^{ax}\ dx = \frac{1}{a}\ e^{ax}$$

and

$$\int xe^{ax}\ dx = \frac{(ax-1)}{a^2}\ e^{ax}$$

$$N = \frac{bu_o}{h^2} + \frac{e^{-ht_1}}{h}\ [g(Q_o + Q_D) - gq_B\ (P_o + \frac{b}{h}) - bgq_B t_1]$$
$$+ \frac{e^{-hT}}{h}\ [-g(Q_o + Q_D - q_B P_o) - b(u_o - gq_B)(T + \frac{1}{h})] \ - \ K$$

However, recall

$$t_1 = \frac{1}{b}\left(\frac{Q_o + Q_D}{q_B} - P_o\right)$$

so that

$$N = \frac{bu_o}{h^2} + \frac{e^{-ht_1}}{h} [g(Q_o + Q_D) - gq_B(P_o + \frac{b}{h}) - g(Q_o + Q_D - P_oq_B)]$$

$$+ \frac{e^{-hT}}{h} [-g(Q_o + Q_D - q_BP_o) - b(u_o - gq_B)(T + \frac{1}{h})] - K$$

$$N = \frac{bu_o}{h^2} - \frac{1}{h} [(bT + \frac{b}{h})(u_o - gq_B) + g(Q_o - q_BP_o)] e^{-hT} \qquad (1)$$

$$-K_o - \frac{1}{h} (ge^{-hT} + jh)Q_D - \frac{gbq_B}{h^2} e^{-\frac{h}{bq_B}(Q_o + Q_D - P_oq_B)}$$

Since all the dependent variables have been eliminated in favor of constants, one can determine the Q_D which will give a maximum benefit stream by taking the derivative of N with respect to Q_D, setting it equal to zero, and solving the resulting equation for Q_D.

Thus

$$\frac{dN}{dQ_D} = 0 = -\frac{1}{h} (ge^{-hT} + jh) - \frac{gbq_B}{h^2} e^{-\frac{h}{b}(\frac{Q_o}{q_B} - P_o)} (-\frac{h}{bq_B}) e^{-\frac{h}{bq_B}Q_D}$$

Therefore

$$ge^{-hT} + jh = ge^{-\frac{h}{b}(\frac{Q_o}{q_B} - P_o)} e^{-\frac{h}{bq_B}Q_D}$$

Taking logarithms of both sides

$$\ln (ge^{-hT} + jh) = \ln g - \frac{h}{b} (\frac{Q_o}{q_B} - P_o) - \frac{h}{bq_B} Q_D$$

$$Q_D = P_oq_B - Q_o + \frac{bq_B}{h} \ln \left\{ \frac{g}{ge^{-hT} + jh} \right\} \qquad (2)$$

Thus if P_o, q_B, Q_o, b, h, g, j, and T are known, the optimal design Q_D can be determined from Equation 2. It is interesting to note that the decision on plant capacity is independent of u_o. Having thus determined Q_D, and knowing u_o, the value of the net benefit stream can be computed from Equation 1. The application of these two equations and the effect of varying the economic parameters g, h, and T are best illustrated by a numerical example.

Numerical Example

P_o = 100,000 Persons

Q_o = 15,000,000 Gallons per day = 15 mgd

q_B = 150 gallons per person per day (This value is based on the average per capita water consumption, including fire demand and industrial uses, in common use today)

i = 2 percent per year

h = 1.98×10^{-2} per year (see Table 1)

b = 1,500 Persons per year (Based on a rate of natural increase of 15 persons per thousand (4))

g = \$0.20 per 1000 gallons (This figure reflects the difference between actual domestic water rates estimated from Table 2 as \$0.35 per 1000 gallons and the actual cost of water treatment estimates at \$15 per million gallons (5))

u_o = \$0.03 per person per day (Estimated from g times q_B)

T = 50 years

j = \$0.10 per gallon per day, capacity (Based on an estimate of cost of a filtration plant of \$100,000 per mgd capacity (5))

K = \$250,000

Using these dimensions, equation 2 becomes

$$Q_D = P_o q_B - Q_o + \frac{bq_B}{h} \ln \left\{ \frac{g}{ge^{-hT} + \frac{jh}{365}} \right\}$$

so that for this example

$$Q_D = \frac{(1.5 \times 10^3)(1.5 \times 10^2)}{(1.98 \times 10^{-2})} \ln \left\{ \frac{2 \times 10^{-4}}{(2\times10^{-4})(e^{-0.99}) + (1\times10^{-1})(5.43\times10^{-5})} \right\}$$

$$Q_D = 10.45 \times 10^6$$

$$Q_D = 10.45 \text{ mgd}$$

Adjusting Equation 1 dimensionally

$$N = \frac{365\ bu_o}{h^2} - \frac{365}{h}\left[\left(\frac{b}{h} + bT\right)(u_o - gq_B) + g(Q_o - q_B P_o)\right] e^{-hT}$$

$$-K_o - \frac{365}{h}\left(ge^{-hT} + \frac{jh}{365}\right) Q_D - 365\ \frac{bq_B g}{h^2}\ e^{-\frac{h}{b}\left(\frac{Q_o}{q_B} - P_o\right)}\ e^{-\frac{h}{bq_B}Q_D}$$

Using the above figures,

$$N = 4.18 \times 10^7 - 1.47Q_D - 4.16 \times 10^7 e^{-0.88 \times 10^{-7} Q_D}$$

The fact that $Q_D = 1.045 \times 10^7$ provides the maximum benefits may be illustrated by plotting a curve of Q_D versus N for the above conditions. Table 3 shows the computation of N for various Q_D's and Figure 4 is a plot of the function.

The dependence of Q_D and N on g can also be illustrated for this simple case. It is assumed for ease of computation that u_o varies with g so that the relation $u_o - gq_B = 0$ is maintained as in the previous example. If all other values remain the same (except u_o and g) then

$$Q_D = 1.135 \times 10^7 \ln \frac{g}{0.37158g + 5.43 \times 10^{-6}}$$

and

$$N = (2.10 \times 10^{11})g - (2.5 \times 10^5)$$

$$- [(6.85 \times 10^3)g + (1 \times 10^{-1})]\ Q_D - (2.08 \times 10^{11})ge^{-(0.88 \times 10^{-7})\ Q_D}$$

Table 4 gives the computation of Q_D for various values of g and Figure 5 shows the functional relation. As is obvious from Figure 5, the dependence of Q_D on g becomes less as g becomes large.

The dependence of N on g has been computed in Table 5 and is shown in Figure 6. As one would infer, N increases almost linearly with g.

The effect of the interest rate and the economic time horizon on the optimal design of the plant is also of interest. These relations are computed in Tables 6 and 7 and shown in Figures 7 and 8. Increasing the interest rate requires decreasing the size of the plant and increasing the economic time horizon requires increasing the plant capacity.

Discussion of Equation 2

As stated earlier, for this simple model the optimum size of the water treatment plant is dependent upon the break-point water use (q_B), the present population (P_o), the present treatment plant capacity (Q_o), the interest rate (as reflected by h), the population growth (b), the cost of treatment plant capacity (j), and the economic time horizon (T). This dependence was represented in Equation 2 as

$$Q_D = P_o \, q_B - Q_o + \frac{bq_B}{h} \ln \left\{ \frac{g}{ge^{-hT} + jh} \right\}$$

The equation can be converted to a more useful form by reducing it to an equation involving dimensionless parameters.

Recall that the time, t_1, at which the total available water per person just equals the break point use, q_B, was determined by

$$t_1 = \frac{Q_D + Q_o - P_o q_B}{bq_B}$$

By substituting the condition for an optimum design as given by Equation 2,

$$t_1 = \frac{1}{h} \ \ln \left\{ \frac{g}{jh + ge^{-hT}} \right\}$$

so that

$$ht_1 = \ln \left\{ \frac{g/jh}{g/jh \ e^{-hT} + 1} \right\} \qquad (3)$$

where ht_1, g/jh, and hT are dimensionless if consistent units are used. Equation 3 can be written more simply by the transformation

$$ht_1 = y$$
$$hT = x$$
$$g/jh = \alpha$$

so that

$$y = \ln \left\{ \frac{\alpha}{\alpha\, e^{-x} + 1} \right\}$$

Note that since j reflects the cost of construction and h the interest rate, the parameter $^{g/}jh$ is a cost-risk factor. That is, high values of $^{g/}jh$ are indicative of high risk and/or low cost situations, whereas low values of $^{g/}jh$ are indicative of low risk and/or high cost situation.

Computations for the solution to Equation 3 are given in Table 8 and the graphical representation is shown in Figure 9.

Conclusions

The economic concepts of utility functions and benefit streams have been applied to the engineering problem of determining the optimal size of municipal water treatment plants. The model presented has been simplified in this preliminary analysis. With the assumptions of a three-parameter utility function, a deterministic population growth, and a linear cost of construction curve, the optimum design is determined by the relationship

$$Q_D = P_o\, q_B - Q_o + \frac{bq_B}{h} \ln \left\{ \frac{g/jh}{g/jh\, e^{-hT} + 1} \right\}$$

The model serves primarily as an introduction to the type of solution proposed. Future models will be designed to include

1. A more realistic four-parameter utility function.
2. A utility function that is dynamic (i.e., reflects the change of q_B with time) rather than static.
3. A population function in which the rate of growth is stochastic rather than deterministic.

4. A cost vector which consists of
    a) An initial cost that depends upon both nominal capacity and provision for flexibility, i.e., the costs of installing equipment that could be overloaded.
    b) Operating costs that depend upon the actual water production during the year, i.e., power, labor, etc.
5. An analysis to include the dynamic case, i.e., timing of construction.
6. An analysis to consider constraints
    a) in budget
    b) in the time horizon within which money must be paid back, such as 15 year bonds, etc.

It is expected that the generalized model can be readily modified to apply to other types of sanitary engineering works such as sewage treatment plants, and can be incorporated in the master computation program for analysis of comprehensive plans for water quality management.

Table 1

$$e^{ht} = (1 + i)^t$$

$$ht \cdot \log_e e = t \cdot \log_e(1+i)$$

$$h = \log_e(1 + i)$$

i	1+i	h
0.010	1.010	0.00995
0.015	1.015	0.01487
0.020	1.020	0.01980
0.025	1.025	0.02468
0.030	1.030	0.02956
0.035	1.035	0.03439
0.040	1.040	0.03922
0.045	1.045	0.04400
0.050	1.050	0.04879

Conversion of Interest Rates

Table 2

Average Domestic Water Rates

Area	Reference	Year	Cost $ per 1000 gallons
Rhode Island	6	1960	0.253-0.320
Los Angeles	7	1960	0.03 - 0.20
Baltimore	7	1960	0.10
Mid Atlantic	8	1960	0.40
General	9	1960	0.30
General (max)	10	1958	0.52

TABLE 3

Determination of N for Various Values of Q_D when

$$N = 4.18 - 1.47\, Q_D - 4.16\, e^{-0.88\, Q_D}$$

Q_D 10^7 GAL/DAY	$0.88\, Q_D$	$e^{-0.88 Q_D}$	$4.16\, e^{-0.88 Q_D}$ I	$1.47\, Q_D$ II	I + II III	N 4.18 - III 10^6
0.1	0.088	0.91577	3.810	0.147	3.957	2.23
0.3	0.264	0.76798	3.190	0.441	3.631	5.49
0.5	0.440	0.64404	2.680	0.735	3.415	7.65
0.8	0.704	0.49461	2.050	1.176	3.226	9.54
1.0	0.880	0.41478	1.723	1.470	3.193	9.87
1.045	0.920	0.39852	1.652	1.535	3.187	9.93
1.1	0.968	0.37984	1.575	1.616	3.191	9.89
1.3	1.144	0.31855	1.325	1.910	3.235	9.45
1.5	1.320	0.26714	1.110	2.205	3.315	8.65
1.8	1.584	0.20516	0.853	2.646	3.499	6.81
2.0	1.760	0.17204	0.715	2.940	3.655	5.25
2.3	2.024	0.13213	0.549	3.381	3.930	2.50
2.5	2.200	0.11080	0.460	3.675	4.135	0.45

TABLE 4

Computation of Q_D for Various Values of g

$$Q_D = \frac{b\ q_B}{h}\ \ell n \left\{ \frac{g}{g e^{-hT} + \frac{\ell h}{365}} \right\}$$

$$Q_D = 1.135 \times 10^7\ \ell n \left\{ \frac{g}{0.37185g + 5.43 \times 10^{-6}} \right\}$$

I g	II 0.37158g	III II + 5.43×10^{-6}	$\frac{I}{III}$	$\ell n \frac{I}{III}$	Q_D 10^6 GAL/DAY
1×10^{-5}	3.7158×10^{-6}	9.1458×10^{-6}	1.09	0.08618	0.98
2×10^{-5}	7.4316×10^{-6}	12.8616×10^{-6}	1.555	0.44147	5.01
1×10^{-4}	3.7158×10^{-5}	4.2588×10^{-5}	2.35	0.85442	9.7
2×10^{-4}	7.4316×10^{-5}	7.9746×10^{-5}	2.51	0.92028	10.45
5×10^{-4}	1.8579×10^{-4}	1.9122×10^{-4}	2.62	0.96317	10.91
8×10^{-4}	2.9726×10^{-4}	3.0269×10^{-4}	2.64	0.97078	11.02
1×10^{-3}	3.7158×10^{-4}	3.7701×10^{-4}	2.65	0.97456	11.05
2×10^{-3}	7.4316×10^{-4}	7.4859×10^{-4}	2.68	0.98582	11.20
4×10^{-3}	1.48632×10^{-3}	1.491×10^{-3}	2.68	0.98582	11.20

TABLE 5

Computation of N for Various Values of g

$$N = (2.10 \times 10^{11})g - (2.5 \times 10^{5}) = [(6.85 \times 10^{3})g + (1 \times 10^{-1})]\, Q_D$$
$$- (2.08 \times 10^{11})g\, e^{-(0.88\times10^{-7})Q_D}$$

$$N = K \quad - \quad Y\ Q_D \quad - \quad X\, e^{-(0.88\times10^{-7})Q_D}$$

$g = 1\times10^{-5}$ $\quad N = 1.85 \times10^{6} - (0.1685)\, Q_D - (2.08\times10^{6})e^{-(0.88\times10^{-7})Q_D}$

$g = 2\times10^{-5}$ $\quad N = 3.95 \times10^{6} - (0.2270)\, Q_D - (4.16\times10^{6})e^{-(0.88\times10^{-7})Q_D}$

$g = 1\times10^{-4}$ $\quad N = 2.075\times10^{7} - (0.785)\, Q_D - (2.08\times10^{7})e^{-(0.88\times10^{-7})Q_D}$

$g = 2\times10^{-4}$ $\quad N = 4.175\times10^{7} - (1.470)\, Q_D - (4.16\times10^{7})e^{-(0.88\times10^{-7})Q_D}$

$g = 5\times10^{-4}$ $\quad N = 1.050\times10^{8} - (3.525)\, Q_D - (1.04\times10^{8})e^{-(0.88\times10^{-7})Q_D}$

$g = 8\times10^{-4}$ $\quad N = 1.658\times10^{8} - (5.580)\, Q_D - (1.664\times10^{8})e^{-(0.88\times10^{-7})Q_D}$

$g = 1\times10^{-3}$ $\quad N = 2.100\times10^{8} - (6.950)\, Q_D - (2.08\times10^{8})e^{-(0.88\times10^{-7})Q_D}$

$g = 2\times10^{-3}$ $\quad N = 4.200\times10^{8} - (13.800)\, Q_D - (4.16\times10^{8})e^{-(0.88\times10^{-7})Q_D}$

g	Q_{max} (from Table 4) 10^6 GAL/DAY	0.88 Q	$e^{-0.88\times10^{-7}Q}$	I X $e^{-0.88Q_D}$ 10^7	II Y Q 10^7	III I + II 10^7	N K - III 10^7 Dollars
1×10^{-5}	0.98	0.0862	0.9174	0.191	0.165	0.356	- 0.171
2×10^{-5}	5.01	0.441	0.6430	0.268	0.118	0.386	+ 0.009
1×10^{-4}	9.7	0.854	0.4257	0.885	0.761	1.646	0.429
2×10^{-4}	10.45	0.920	0.3985	1.655	1.535	3.187	0.988
5×10^{-4}	10.91	0.960	0.3829	3.98	3.85	7.83	2.67
8×10^{-4}	11.02	0.970	0.3791	6.31	6.05	12.36	4.22
1×10^{-3}	11.05	0.972	0.3783	7.88	7.68	15.56	5.44
2×10^{-3}	11.20	0.985	0.3734	15.50	15.45	30.95	11.05

TABLE 6

Computation of Design Capacity for Various Values of Interest Rate

$$Q_D = \frac{b\ q_B}{h}\ \ell n \left\{ \frac{g}{ge^{-hT} + \frac{jh}{365}} \right\}$$

i	$-hT$	e^{-hT}	ge^{-hT} 10^{-4}	$\frac{jh}{365}$ 10^{-4}	$ge^{-hT} + \frac{jh}{365}$	$\frac{g}{ge^{-hT} + \frac{jh}{365}}$	$\ell n \left\{ \right\}$	$\frac{b\ q_B}{h}$ 10^6	Q_D 10^6
1	0.4975	0.60809	1.21618	0.0273	1.2435	1.605	0.47312	22.60	10.65
2	0.990	0.37158	0.74316	0.0543	0.7875	2.51	0.92028	11.35	10.45
3	1.48	0.22764	0.45528	0.0811	0.5364	3.73	1.31641	7.60	10.35
4	1.96	0.14086	0.28172	0.1072	0.38892	5.14	1.63705	5.74	9.40
5	2.44	0.08716	0.17432	0.1340	0.30832	6.49	1.87026	4.62	8.65

TABLE 7

Computation of Design Capacity for Various Values of T

$$Q_D = \frac{b\ q_B}{h}\,\ell n \left\{ \frac{g}{g e^{-hT} + \frac{ih}{365}} \right\}$$

$$= (11.35 \times 10^6)\ \ell n \left\{ \frac{20 \times 10^{-5}}{20\times10^{-5}\ e^{-hT} + 0.543\times10^{-5}} \right\}$$

T	hT	e^{-hT}	I $20\ e^{-hT}$	II I+0.543	III 20/II	IV ℓn III	Q_D 10^6 GAL/DAY
0	0	1	20	20.543	0.976	-0.024	-0.272
25	0.495	0.60958	12.1916	12.7346	1.57	+0.45108	+5.11
50	0.99	0.37158	7.4316	7.9746	2.51	0.92028	10.45
75	1.485	0.22650	4.5300	5.0730	3.94	1.37118	15.6
100	1.98	0.13807	2.7614	3.3044	6.06	1.80171	20.4
120	2.38	0.09255	1.8510	2.3940	8.38	2.12585	24.1
140	2.77	0.06266	1.2532	1.8062	11.10	2.30695	26.2
150	2.97	0.05130	1.0260	1.5690	12.75	2.44553	27.7
160	3.17	0.04202	0.8404	1.3834	14.5	2.67415	30.3
180	3.56	0.02844	0.5688	1.1118	17.9	2.88481	32.7
200	3.96	0.01906	0.3812	0.9242	21.7	3.07732	34.9
300	5.94	0.00253	0.0506	0.5936	33.8	3.52047	40.0
400	7.92	0.00037	0.0074	0.5504	36.3	3.59182	40.8

TABLE 8

Computation of $y = \ln \dfrac{\alpha}{1 + \alpha\, e^{-x}}$

α	x	e^{-x}	$\alpha\, e^{-x}$	$1 + \alpha e^{-x}$	$\dfrac{\alpha}{1 + \alpha e^{-x}}$	$y = \ln \dfrac{\alpha}{1 + \alpha e^{-x}}$
α= 50	0.3	0.741	37.1	38.1	1.31	0.270
	0.5	0.607	30.4	31.4	1.59	0.464
	1.0	0.368	18.4	19.4	2.58	0.948
	1.5	0.223	11.2	12.2	4.10	1.411
	2.0	0.135	6.75	7.75	6.45	1.864
	2.5	0.0821	4.11	5.11	9.79	2.281
	3.0	0.0498	2.49	3.49	14.3	2.660
	3.5	0.0302	1.51	2.51	19.9	2.991
	4.0	0.0183	0.915	1.92	26.1	3.262
α= 16	0.3	0.741	11.85	12.85	1.24	0.215
	0.5	0.607	9.72	10.72	1.49	0.399
	1.0	0.368	5.89	6.89	2.32	0.842
	1.5	0.223	3.57	4.57	3.52	1.26
	2.0	0.135	2.16	3.16	5.05	1.62
	2.5	0.0821	1.31	2.31	6.93	1.94
	3.0	0.0498	0.796	1.80	8.88	2.18
	3.5	0.0302	0.483	1.48	10.8	2.38
	4.0	0.0183	0.293	1.29	12.4	2.52
α= 8	0.3	0.741	5.92	6.92	1.15	0.140
	0.5	0.607	4.86	5.86	1.36	0.307
	1.0	0.368	2.94	3.94	2.03	0.708
	1.5	0.223	1.78	2.78	2.88	1.058
	2.0	0.135	1.08	2.08	3.85	1.348
	2.5	0.0821	0.655	1.655	4.83	1.575
	3.0	0.0498	0.398	1.398	5.72	1.744
	3.5	0.0302	0.241	1.241	6.44	1.863
	4.0	0.0183	0.146	1.146	7.00	1.946
α= 4	0.3	0.741	2.96	3.96	1.01	0.00995
	0.5	0.607	2.43	3.43	1.16	0.148
	1.0	0.368	1.47	2.47	1.62	0.482
	1.5	0.223	0.892	1.892	2.12	0.751
	2.0	0.135	0.541	1.541	2.60	0.956
	2.5	0.0821	0.328	1.328	3.01	1.102
	3.0	0.0498	0.199	1.199	3.33	1.203
	3.5	0.0302	0.121	1.121	3.57	1.273
	4.0	0.0183	0.0733	1.073	3.74	1.319

TABLE 8 (continued)

$\alpha = 2$	0.3	0.741	1.48	2.48	0.807	-0.214
	0.5	0.607	1.21	2.21	0.905	-0.099
	1.0	0.368	0.736	1.736	1.15	+0.140
	1.5	0.223	0.446	1.446	1.38	0.322
	2.0	0.135	0.270	1.270	1.57	0.451
	2.5	0.0821	0.164	1.164	1.72	0.542
	3.0	0.0498	0.0996	1.100	1.82	0.599
	3.5	0.0302	0.0604	1.060	1.89	0.637
	4.0	0.0183	0.0366	1.037	1.92	0.652
$\alpha = 1$	0.3	0.741	0.741	1.74	0.575	-0.553
	0.5	0.607	0.607	1.61	0.620	-0.478
	1.0	0.368	0.368	1.37	0.730	-0.315
	1.5	0.223	0.223	1.22	0.820	-0.198
	2.0	0.135	0.135	1.14	0.877	-0.131
	2.5	0.0821	0.0821	1.08	0.925	-0.078
	3.0	0.0498	0.0498	1.05	0.952	-0.049
	3.5	0.0302	0.0302	1.03	0.971	-0.029
	4.0	0.0183	0.0183	1.02	0.980	-0.020

References

1. Fair, Gordon M. and Geyer, John C., *Water Supply and Waste-Water Disposal*, John Wiley & Sons, Inc., New York, N.Y., 1954, pp. 116-118.

2. Turneaure, F.E. and Russell, H.L., *Public Water Supplies*, John Wiley & Sons, Inc., New York, N. Y., 1940.

3. Thomas, Harold A. and Watermeyer, Peter, in Chapter 14 of Maass,Arthur, et al, *Design of Water Resource Systems*, Harvard University Press, Cambridge, Mass., 1962.

4. Fair and Geyer, p. 119.

5. Fair and Geyer, p. 53.

6. Holton, Phillip J., Jr., "Water Rates Go Up after 36 Years," *Water Works Engineering*, *113*, p. 112 (Feb. 1960).

7. "Water Reuse Beats Desalting," *Engineering News Record*, *165*, 18, p.24 (Nov. 3, 1960).

8. "Desalting the Sea," *The Wall Street Journal*, Thursday, Aug. 11, 1960.

9. Dodge, Barnett F., "Fresh Water from Saline Waters - An Engineering Research Problem," *American Scientist*, *48*, 476-514 (Dec. 1960).

10. Koenig, Louis, "Economic Boundaries of Saline Water Conversion," *J. Am. Water Works Assoc.*, *51*, 845 (1959).

5-22

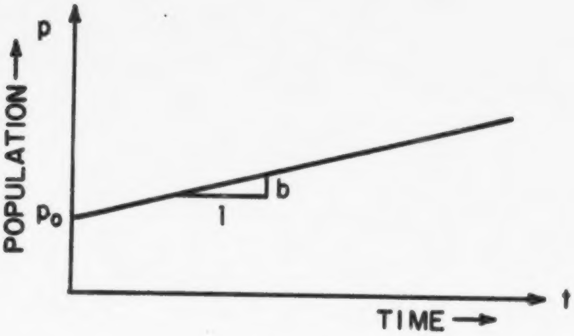


FIG. 1 POPULATION FUNCTION

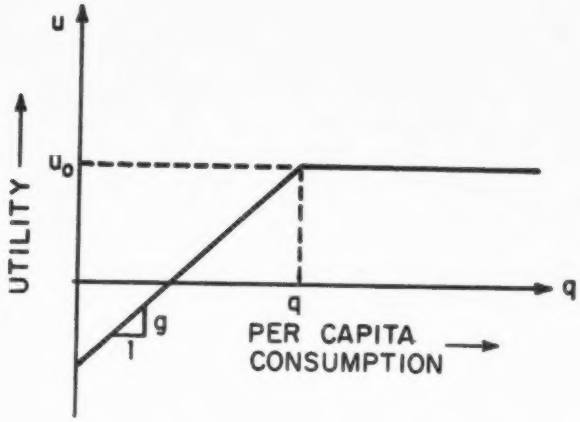


FIG. 2 UTILITY FUNCTION

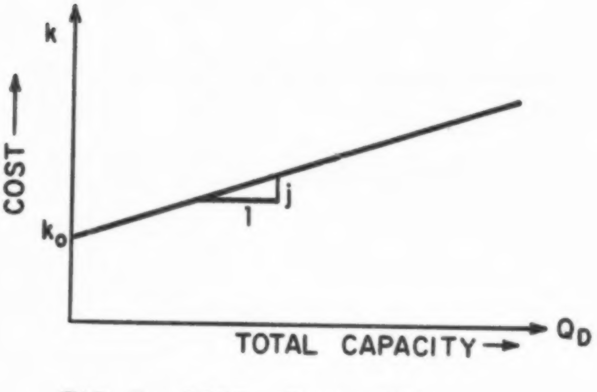
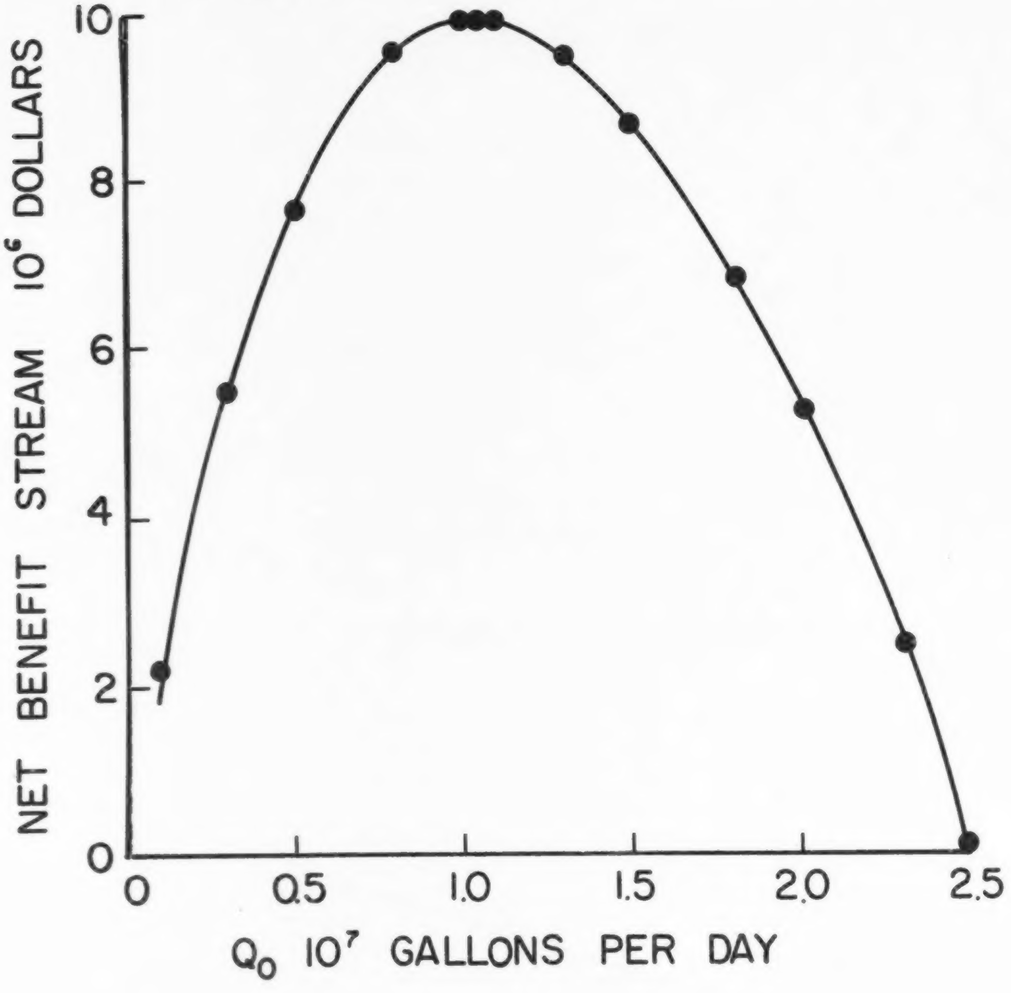


FIG. 3 COST FUNCTION

5-23

Figure 4

DESIGN CAPACITY- NET BENEFITS



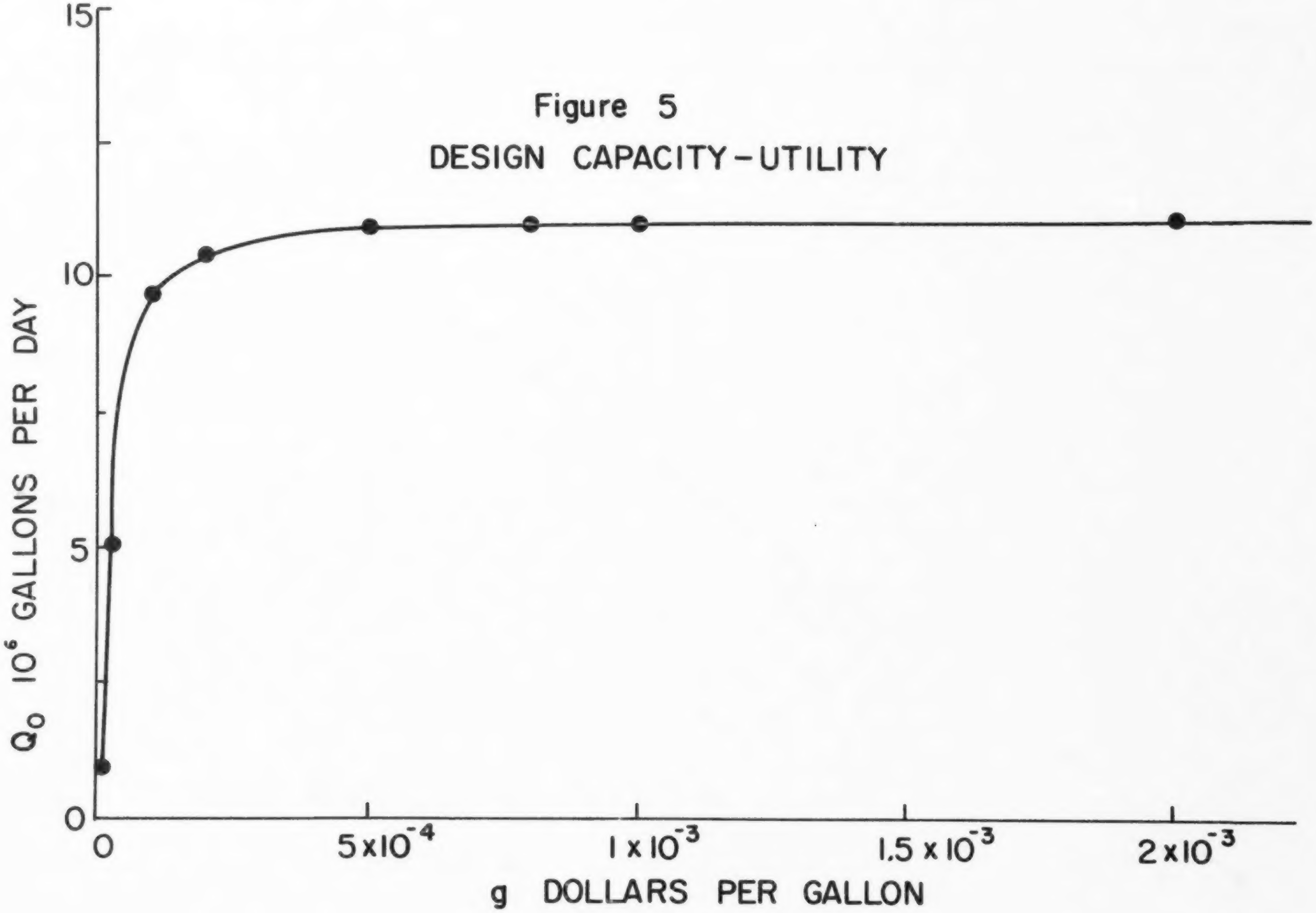


Figure 5

DESIGN CAPACITY-UTILITY

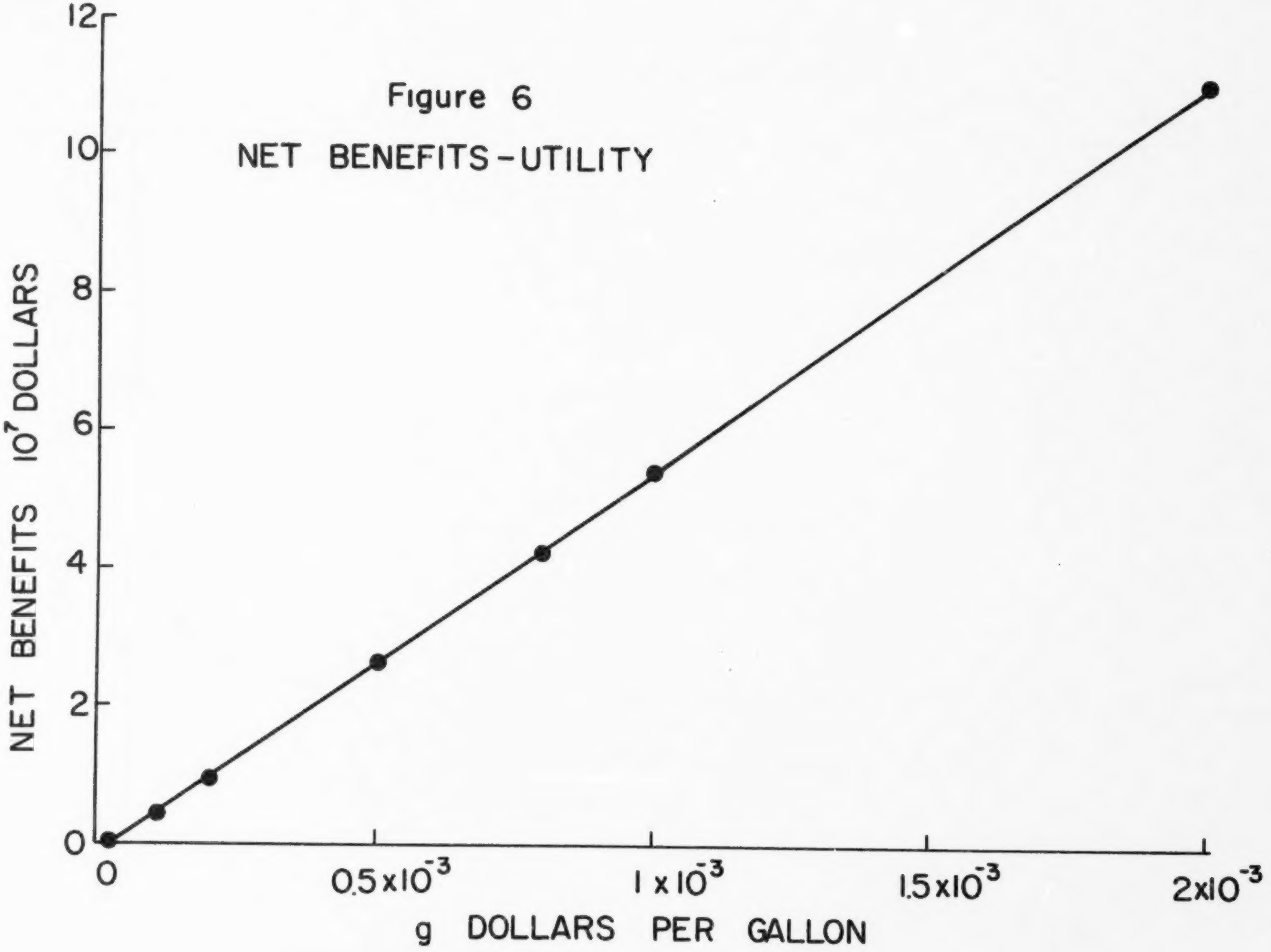
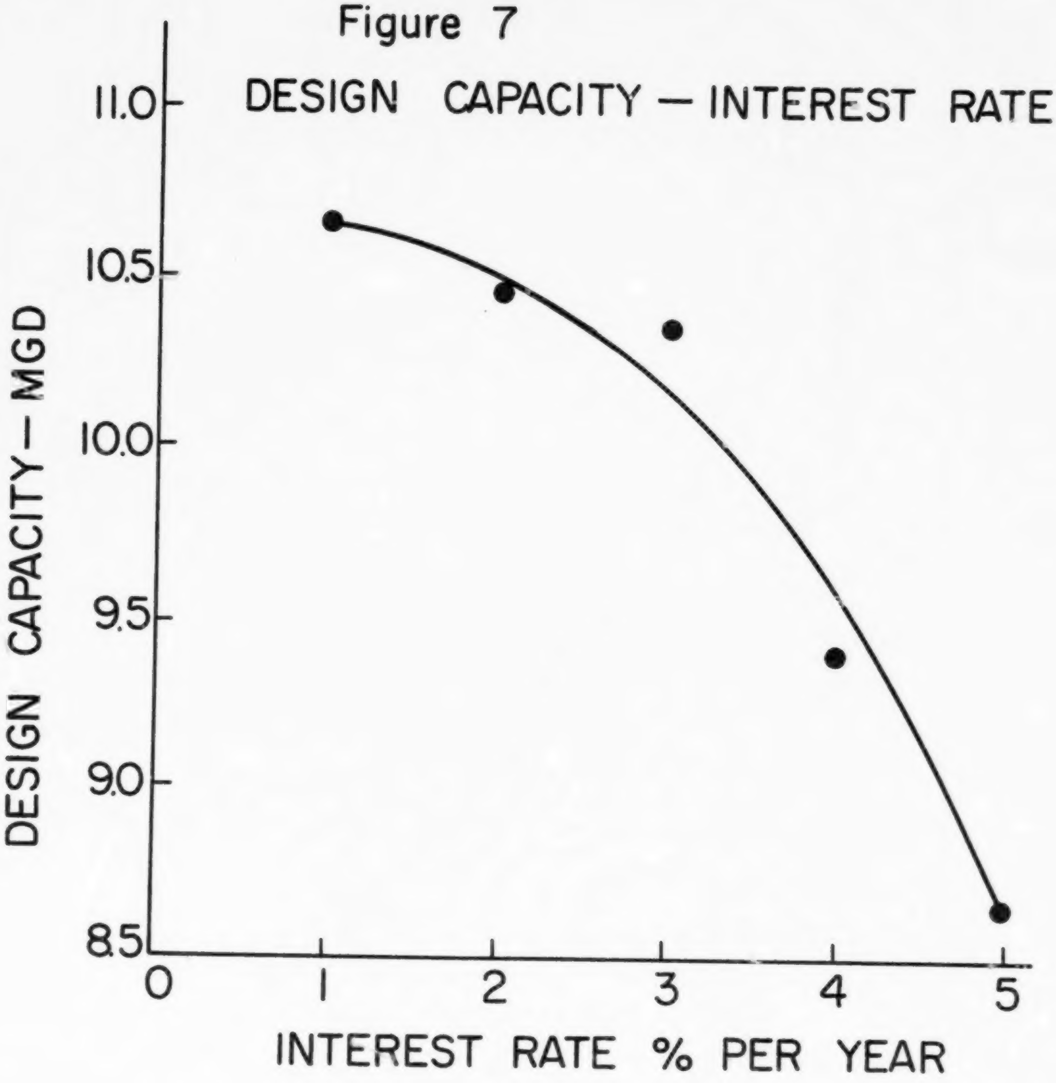


Figure 6

NET BENEFITS-UTILITY

Figure 7

DESIGN CAPACITY — INTEREST RATE



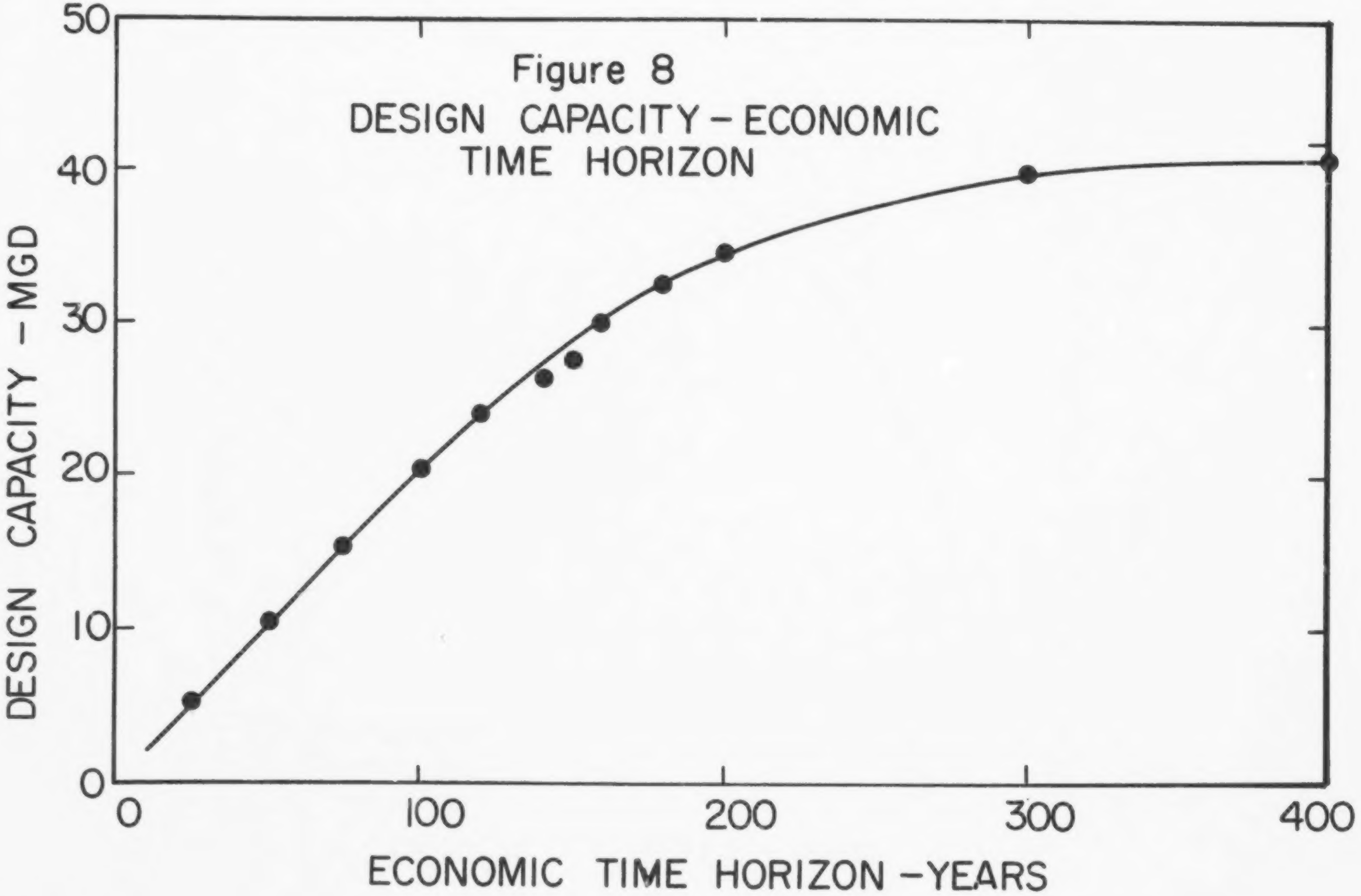


Figure 8
DESIGN CAPACITY - ECONOMIC
TIME HORIZON

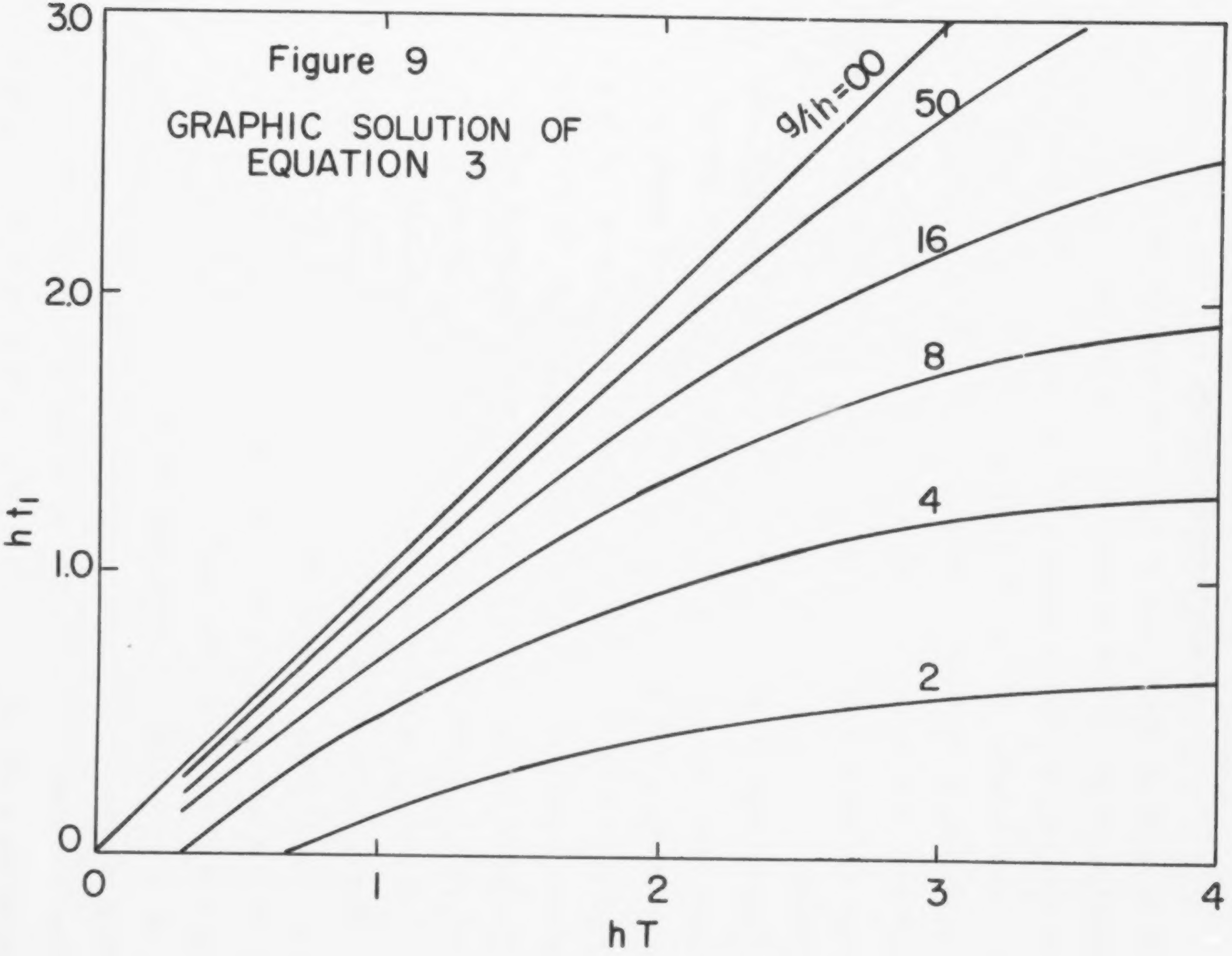


Figure 9

GRAPHIC SOLUTION OF EQUATION 3

5-29

CHAPTER SIX

A Method of Optimizing Designs of Water Quality Management Systems

To carry out systems analysis of problems of water quality, complex mathematical formulations must be developed to simulate nature. The set of results that satisfy both the imposed conditions and the given conditions may comprise many answers. To identify the optimal set of design variables, or the choice resulting in the decision that most closely achieves the plan objectives, requires specialized analytical methods. In this chapter a generalized technique for solving problems of this type has been developed and adapted to machine computation.

In developing and applying the theory two examples are discussed. The first is a simple example with only two independent variables so that the principle of the method can be shown graphically. The second example is a practical application - the problem is to find the dimensions of a rectangular sedimentation tank which will provide a specified detention time for a given flow at minimum cost with a restriction on the ratio of length to width of the basin. In this example the independent variables are the three dimensions of the tank. The cost, a function of these three variables, is to be minimized. The constraints are that the volume of the tank be not less than required by the specified flow and detention time and that the length of the tank be not less than a given multiple of its width. However, the method developed is designed to cope with much more complex problems.

In the case of only two independent variables, x_1 and x_2, the given function of these variables takes on different values and defines a "response surface." As shown in Figure 1, this may be represented in two dimensions as a contour map of the response surface. Pairs of values of x_1 and x_2 lying on the same contour line represent different designs having the same cost. This surface may have one or many hills and valleys. The problem is to determine which of the points on the map is highest. Constraints between the design variable limit the search to certain regions of the map (see Figure 2).

The procedure used is a so-called method of "steepest ascent."* It can be visualized as the activity of a group of paratroopers walking up the closest hill. The paratroopers are dropped at random over the constrained region of the response surface. Each has orders to start climbing the steepest slope from the point at which he lands. However, some of the paratroopers will not be able to walk continuously up the steepest slope because, as shown in Figure 2, they will run into the constraints which act as a solid wall to stop their progress. Upon hitting the wall, the paratroopers rebound a few steps - always in an uphill direction. In this blundering but systematic movement the paratroopers eventually reach the highest point available to them and report its altitude to the command headquarters. From this information the highest peak corresponding to the optimal design is identified.

If we have a response surface in n + 1 space with n independent variables, represented by

$$B = f\,(x_1, x_2, \ldots\ldots\ldots x_n)$$

then

$$dB = \frac{\partial B}{\partial x_1}\, dx_1 + \frac{\partial B}{\partial x_2}\, dx_2 \cdots\cdots \frac{\partial B}{\partial x_n}\, dx_n \qquad (1)$$

if we evaluate the response surface at some point $x_1^0, x_2^0, \ldots\ldots x_n^0$ then $B = B^0$. We now move some finite distance in each direction (for convenience only two variables are used but this proof holds for the n variables), let these distances be u, v, etc. Then the new point, call it $x_1^1, x_2^1, x_3^1 \ldots\ldots$, will lie on the circle shown in Figure 1. The value of B at this point will be B^1. Although u, v, ... are finite, they are small, therefore approximately

---

* For problems of identifying the minimum cost design the procedure might more aptly be called the method of "steepest descent." However, in more general applications both costs and benefits are under consideration and the objective is to identify the design that maximizes net benefits and so the search for the optimal set of parameters entails an ascent of the highest peak.

$$dB = B^1 - B^0 .$$

Also let the distance moved from point 1 to point 2 be ρ. Then

$$\rho^2 = u^2 + v^2 \ldots\ldots z^2 \tag{2}$$

The object is now to maximize dB subject to the constraint equation (2).

Considering the three dimensional case and making use of Lagrange's Method of Undetermined Multipliers we can write equation (1) as

$$\Psi = \frac{\partial B}{\partial x_1} u + \frac{\partial B}{\partial x_2} v$$

and equation (2) as

$$\theta = u^2 + v^2 - \rho^2 = 0. \tag{3}$$

Differentiating with respect to each of the variables and using λ as the Lagrangean Multiplier we have

$$\frac{\partial \Psi}{\partial u} + \lambda \frac{\partial \theta}{\partial u} = 0$$

$$\frac{\partial \Psi}{\partial v} + \lambda \frac{\partial \theta}{\partial v} = 0$$

gives

$$\frac{\partial B}{\partial x_1} - \lambda 2u = 0 \tag{4}$$

$$\frac{\partial B}{\partial x_2} - \lambda 2v = 0 \tag{5}$$

solve for λ

$$\lambda = \frac{\partial B}{\partial x_1} / 2u = \frac{\partial B}{\partial x_2} / 2v$$

therefore,

$$u = \frac{\partial B}{\partial x_1} / 2\lambda \qquad v = \frac{\partial B}{\partial x_2} / 2\lambda \tag{6}$$

substituting in the constraint equation (3)

$$\lambda = \frac{1}{2\rho}\sqrt{\left(\frac{\partial B}{\partial x_1}\right)^2 + \left(\frac{\partial B}{\partial x_2}\right)^2}$$

substituting in equation (6)

$$\begin{gathered} u = \frac{\frac{\partial B}{\partial x_1}\ \rho}{\sqrt{\left(\frac{\partial B}{\partial x_1}\right)^2 + \left(\frac{\partial B}{\partial x_x}\right)^2}} \\ \\ v = \frac{\frac{\partial B}{\partial x_2}\ \rho}{\sqrt{\left(\frac{\partial B}{\partial x_1}\right)^2 + \left(\frac{\partial B}{\partial x_2}\right)^2}} \end{gathered} \tag{7}$$

equations (7) can now be generalized, in n dimensions, thus

$$v = \frac{\frac{\partial B}{\partial x_1}\ \rho}{\sqrt{\left(\frac{\partial B}{\partial x_1}\right)^2 + \left(\frac{\partial B}{\partial x_2}\right)^2 + \ldots.\left(\frac{\partial B}{\partial x_n}\right)^2}}$$

$$v = \frac{\frac{\partial B}{\partial x_2}\ \rho}{\sqrt{\left(\frac{\partial B}{\partial x_1}\right)^2 + \left(\frac{\partial B}{\partial x_2}\right)^2 + \ldots\left(\frac{\partial B}{\partial x_n}\right)^2}}$$

$$\vdots$$

$$z = \frac{\frac{\partial B}{\partial x_n}\ \rho}{\sqrt{\left(\frac{\partial B}{\partial x_1}\right)^2 + \left(\frac{\partial B}{\partial x_2}\right)^2 + \ldots\left(\frac{\partial B}{\partial x_n}\right)^2}}$$

We now know the exact location of x^1 on the circle. It is at

$$x_1^{\ 1} = x_1^{\ 0} + u$$

$$x_2^{\ 1} = x_2^{\ 0} + v$$

$$\vdots$$

$$x_n^{\ 1} = x_n^{\ 0} + z$$

To actually compute the values of u, v, ... z we need only the equation of the response surface and the size of the step to be taken, ρ. The choise of ρ is a crucial point in the steepest ascent method and is discussed later. However, assuming for one moment that we know ρ, or as we refer to it in the programs as STEP, we can compute the u, v, . . . z and move to the new point x^1 where the process is then repeated. The path taken to the maximum is shown on the diagram Figure 1.

By slightly modifying the constraints, we have adapted the unconstrained method to the constrained case with inequalities as constraints.

Selection of Step Size

Selection of the step size is of critical importance in steepest ascent methods and in fact for all gradient methods. Crockett and Chernoff[1] discuss the selection of the step size in terms of its relationship to the eigenvalues of the matrix of second partial derivatives. Zoutendijk[2] also outlines a method of determining the size of the step. They are concerned with a reappraisal of the step size after each step. While we realize this to be very important our models do not incorporate such a device. Here, we have started out with an outright guess at the size of the step, STEP. However, we did incorporate devices to reduce STEP to STEP/2.0 when we approach the peaks and/or constraints. As this is an iterative solution, we have the size of step decreasing at a geometric rate of one-half at these points.

---

[1] Crockett, J. B. and H. Chernoff, "Gradient Methods of Maximization" Pacific J. Math. 5, 33 (1955).

[2] Zoutendijk, G. "Method of Feasible Direction" Elsevier Pub. Co Amsterdam.

Sampling Procedure

The other very critical factor in the steepest ascent method of analysis of response surfaces is the method used to ensure reaching the highest peak. We have chosen to use the random sampling method outlined in Chapter 10 in the "Design of Water Resource Systems."[3] This method does not ensure a given probability of reaching the highest peak. What it does give is a probability of the random sample covering the population of the peaks. However, if the response surface is jagged the peak chosen may be far below the optimum.

In the absence of any knowledge of the shape of the surface the use of systematic sampling is ruled out because of the vast number of grid points required to adequately cover the area. With two points on each variable and 2 variables we require $2^2 = 4$ grid points for three variables, $2^3 = 8$, for 10 variables 2^{10}.

Didactic Model Number One

In order to have a model which would be easy to follow and check, we chose a simple three dimensional surface with two peaks of known altitude and location (this is shown in Figure 3). We chose three constraints

$$x \geq 1.5$$

$$y \geq 0$$

$$x + y \leq 20$$

---

[3] Maas, Hufschmidt, Dorfman, Thomas, Marglin and Fair, "Design of Water Resource Systems" Harvard University Press, Cambridge, Mass., 1962.

The equation of the surface was

$$z = \frac{K}{s + (x-a)^2 + (y-b)^2} + \frac{L}{t + (x-c) + (y-d)^2}$$

which had maxima at (a,b) and (c,d), K and L are measures of their relative heights and s and t are indicators of their slopes. For simplicity we chose

$s = t = 1$

$K = 10$

$L = 20$

(a,b) as (2,2) inside the constrained region

(c,d) as (10,12) outside the constrained region

Hence,

$$z = \frac{10}{1 + (x-2)^2 + (y-2)^2} + \frac{20}{1 + (x-10)^2 + (y-12)^2}$$

which has a maximum of $z = 20.121$ at (10,12)

and $z = 10.121$ at (2,2).

By simple geometry the height of the shoulder was found to be 7.03 and its location at (9.11, 10.89). This is shown in Figure 4.

The program for this model was as shown in Figure 5. The variables used were X(1) for x and X(2) for y.

Three checking segments were incorporated in the program. The first, executed at the start of the main routine, examines the random variables generated at the start of each sample in order to ensure that the initial point lay in the constrained region. The second segment, also in the main routine, verifies that all the partial derivatives are zero or within a specified tolerance. This segment includes a branch for determining if the peak was overstepped, and if so,for returning to the previous portion and reducing the step size. The third checking segment, which comprises subroutine CHECK, is shown in Figure 6; Figure 7 is a graphical representation of the logic of subroutine CHECK.

Economic Dimensions of Rectangular Sedimentation Basins

Application of the steepest ascent method runs into complications when a system involving mixed units is involved. The method outlined in Section II assumes that the increments, u and v, are stated in the same units, as was the case with the didactic model. To overcome this difficulty we have made the units of the system dimensionless by stating them as ratios of their permissible range.

Hence X(J) expressed in natural units can be converted into X(J) expressed as a percentage-range ratio as follows:

X(J) = XMIN(J) + X(J)(as a range unit)·(XMAX(J)-XMIN(J))/100

where XMAX(J) and XMIN(J) are the maximum and minimum values of X(J) in natural units, and similarly X(J) in range units,

X(J) = (X(J)(as a natural unit)-XMIN(J))·100/(MAX(J)-XMIN(J)

We are required to find the least cost dimensions of a rectangular sedimentation tank given the following conditions,

Retention period = 2 hours
Flow = 2 mgd.
a = cost of roof per unit area, in dollars per square foot
c = cost of floor per unit area, in dollars per square foot
ℓ = length of basin, in feet
w = width of basin, in feet
h = depth of basin, in feet
kh^n = cost of outer wall of basin, in dollars per foot
k = 1.55
n = 1.5
v = volume of basin in cubic feet

given

a = 2c = 6 dollars per square foot

Total cost to be minimized

$$C_T = a \ell w + c \ell w + kh^n (2\ell + 2w)$$

subject to the constraints

$$w \leq \ell \cdot f.$$

where

$$f = \frac{1}{4}$$

and that

$$h \ell w \geq v$$

For the model we are investigating the program was written using the following:

c called COST = 3 dollars per square foot

n called EN = 1.5

k called CAY = 1.55

f called F = 0.25

h called X(1) XMIN(1) = 5.0 ft. XMAX(1) = 30.0 ft.

ℓ called X(2) XMIN(2) = 50.0 ft. XMAX(2) = 100.0 ft.

w called X(3) XMIN(3) = 10.0 ft. XMAX(3) = 25.0 ft.

v called VOL = 22,300 cu. ft.

Number of variables called N = 3.

ATOL = tolerance on volume = 100 cu. ft.

BTOL = tolerance on width = 0.25 ft.

CTOL = tolerance on cost = 20 dollars

DTOL = tolerance on slope = 0.0001

ETOL = tolerance on amount of rebound at constraint = 0.300

TOL = check for minimum = 0.010

NRAND = number of random samples = 5

IRAND = trigger for random samples = 1

RPAT = trigger for random samples = 2.0

STEPA = ρ, the step size = 1.0

XMINR = minimum range value of X = 0.0

XMAXR = maximum range value of X = 100.0

DX(J) = increment used to find partial derivatives = 0.010

D(J) = slope in the J direction

ICHK = control to restrict the number of iterations = 10

KOOP = number of samples

KTOL = number of steps at constraint ± tolerance

ICON = number of steps at constraint

ITOL = internal counter

KCHK = internal counter

The results of the didactic model were as follows:

1st random sample	1 KOOPS maximum at	8.92176,	11.07360	value	7.3569
2nd random sample	2 KOOPS maximum at	9.04464,	10.94756	value	7.0186
3rd random sample	3 KOOPS maximum at	10.40274,	9.59464	value	3.0591
4th random sample	4 KOOPS maximum at	1.99653	1.99506	value	10.121
5th random sample	5 KOOPS maximum at	8.95841,	11.03757	value	7.3808

On examining these results we can see that the third sample did not attain the constrained maximum. This is due to the fact that 3 KOOPS started a long way from the maximum and ICON exceeded ICHK before it reached the maximum. ICHK was put in to control the amount of time the computer would spend in attaining each maxima. If ICHK were made larger, say 99, then the 3 KOOPS sample would have reached the constrained peak. The discrepancies between 1 KOOPS, 2 KOOPS, and 5 KOOPS were due to the size of ATOL chosen and could be made much smaller by choice of a smaller ATOL. However, even under these conditions, of the 5 samples taken 4 of them reached the peaks of the surface. Also each of the peaks of the surface was covered by the random sampling.

The results of the rectangular sedimentation basin problem are listed.

	Cost dollars	Height feet	Length feet	Width feet	Volume cubic feet
1 KOOPS	32,377	18.8206	68.3333	17.3730	22,342
2 KOOPS	41,771	22.5119	88.5075	11.0538	22,024
3 KOOPS	32,832	16.7949	83.5389	15.5483	21,814
4 KOOPS	32,754	13.3651	85.2855	21.5838	24,602
5 KOOPS	41,170	29.2433	64.3223	12. 5683	22,024

Under these conditions, we see that the optimal agreement would be to use the 1 KOOPS sample. This gives the least cost of $32,377 for the basin.

Changing ICHK to 20 and selecting a set of new random starting points gives:

	Cost	Height	Length	Width	Volume
1 KOOPS	30,693	10.8069	93.4395	21.3949	21,604
2 KOOPS	36,761	18.7734	91.2847	12.7986	21,933
3 KOOPS	31,455	10.0079	98.9620	21.9881	21,770
4 KOOPS	32,183	14.1282	94.9800	16.2312	21,780
5 KOOPS	30,769	10.8419	93.5866	21.4195	21,733

The least cost is $30,693 at 10.8069, 93.4395, and 21.3949. The discrepancies between the volume and the required volume is due to the use of tolerances on the descent at the constraints. These discrepancies could be reduced by reducing the tolerances and increasing ICHK

Solutions by Lagrangian Method of Undetermined Multipliers

A solution was made by hand in which the constraints were taken to be equalities.

$$C_T = a \ell w + c \ell w + kh^n (2 \ell + 2w)$$

$$C_T = 3 c \ell w + 2 kh^n (\ell + w)$$

where

$$w = \ell\ f$$

$$C_T = 3\ c\ f\ \ell^2 + 2\ kh^n\ \ell\ (\ 1 + f)$$

$$\emptyset = 3\ c\ \ell^2\ f + 2\ kh^n\ \ell\ (1 + f) - \lambda\ (h\ \ell^2 f - v)$$

$$\frac{\partial \emptyset}{\partial \ell} = 6\ c\ \ell\ f + 2\ kh^n\ (1 + f) - 2\ \lambda\ h\ \ell\ f = 0 \qquad (1)$$

$$\frac{\partial \emptyset}{\partial \ell} = \qquad 2n\ kh^{n-1}\ \ell\ (1 + f) - \lambda\ \ell^2\ f = 0 \qquad (2)$$

$$\frac{\partial \emptyset}{\partial \lambda} \qquad h\ \ell^2\ f - v = 0 \qquad (3)$$

from (2) $\lambda = \dfrac{2\ n\ kh^{n-1}(1 + f)}{\ell f}$

with $\quad \ell = (\frac{v}{hf})^{1/2}$

$$h^2 = \frac{9v^{1/2}}{1.55x5} = 1.16v^{1/2} = 17.4$$

$$h = 13.169 \text{ ft.}$$

$$\ell = \frac{v^{1/2}}{h^{1/2}\ f^{\ 1/2}} = \frac{150\ x\ 2}{3.67\ x\ 1}$$

$$= 82.301 \text{ ft.}$$

$$w = 20.575 \text{ ft.}$$

$$h = 13.169 \text{ ft.}$$

$$\ell = 82.301 \text{ ft.}$$

$$w = 20.575 \text{ ft.}$$

$v = 22{,}299.6$ cu. ft. $\qquad C_T = \$30{,}480$

check second derivatives

$$\frac{\partial^2 \emptyset}{\partial \ell^2} = 6\ c\ f - 2\lambda hf \geq 0 \qquad (A)$$

$$\frac{\partial^2 \emptyset}{\partial h^2} = 2n(n-1)\ \ kh^{n-2}\ \ \ell\ (1 + f)\ \ \geq 0 \qquad (B)$$

Condition (B) always satisfied

Condition (A) is satisfied when

$$3c - \lambda h \geq 0$$

$$9 - 13.2\,\lambda \geq 0$$

$$\frac{9}{13.2} - \lambda \geq 0$$

$$.68 - \lambda \geq 0$$

$\therefore$ to satisfy condition (A) $\lambda \leq .68$

$$\text{now } \lambda = \frac{2n\, kh^{n-1}\, \ell\,(1+f)}{\ell^2 f} = \frac{2 \times 1.5 \times 1.55 \times 13.2^{1/2}}{81.6 \times 1/4} \quad \frac{5}{4}$$

$$= \frac{4.65 \times 3.67 \times 5}{81.6} = 1.045 \not\leq .68$$

Hence this is not a minimum but a saddle point.

To satisfy condition (A)

$$3c - \lambda h \geq 0$$

$$\frac{3c}{h} - \lambda \geq 0$$

$$\therefore \quad \lambda \leq \frac{3c}{h}$$

$$\therefore \quad \frac{2n\, kh^{n-1}(1+f)}{\ell f} \leq \frac{3c}{h}$$

$$\ell \geq 2.59\, h^n$$

e.g.

$$\left.\begin{array}{l} \text{If } h = 10 \\ \text{then } \ell = 83 \\ \text{and } w = 20.75 \end{array}\right\} \quad \text{Vol} = 17{,}500 \text{ cu. ft.}$$

and c = \$25,655

It appears that, subject to the constraints, condition (A) can never be fulfilled.

Hence the saddle point

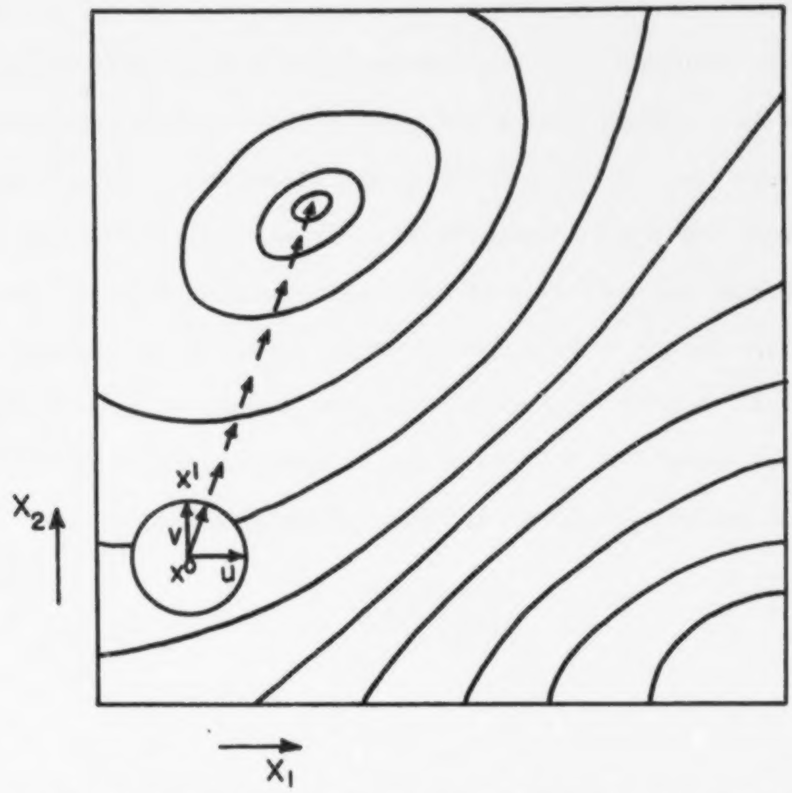
$$\left.\begin{aligned} h &= 13.169 \\ \ell &= 82.301 \\ w &= 20.575 \end{aligned}\right\} \quad v = 22{,}299.6 \text{ cu. ft.}$$

is the lowest point of the response surface at the constraints.

From the results of the hand computation we see that the least costs are $30,480 by hand and $30,693 by machine at 13.169, 82.301, 20.575 and at 10.8069, 93,4395, 21.3949, respectively. The discrepancies can be accounted for by the fact that there is a plateau on the response surface over this area and that the cost is indifferent to a wide variation in the dimensions in this region. The difference between the machine computation of least cost and the hand computation lies in current insensitivity of the machine method. The machine method could be made more sensitive by a more suitable choice of the various tolerances and steps assumed. However, the choice of these depend, to a large extent, upon experience. This experience will be gained by the application of the method to various types of economic-engineering optimization problems.

6-14

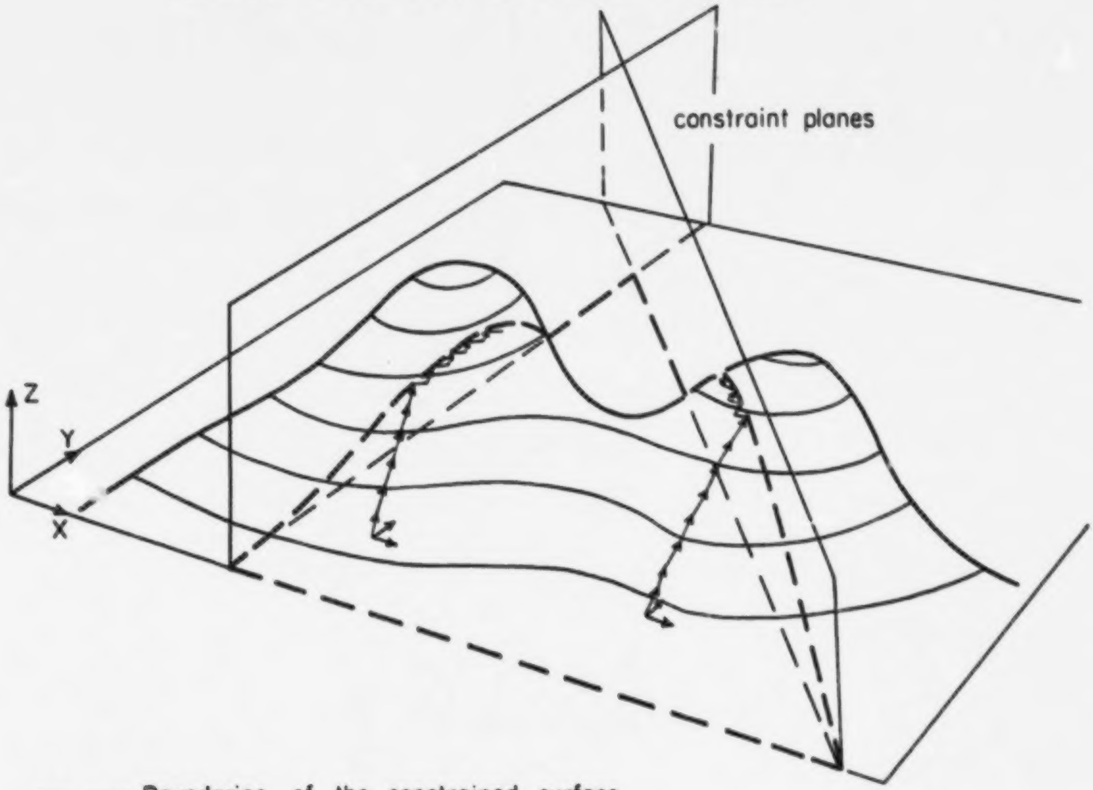
Figure 1



This is a two dimensional representation of the three dimensional surface

Figure 2

Paratroopers Walking Up the Constrained Surface


constraint planes
Z
Y
X

— — Boundaries of the constrained surface.

Figure 3

Shape of Surface for the Didactic Model

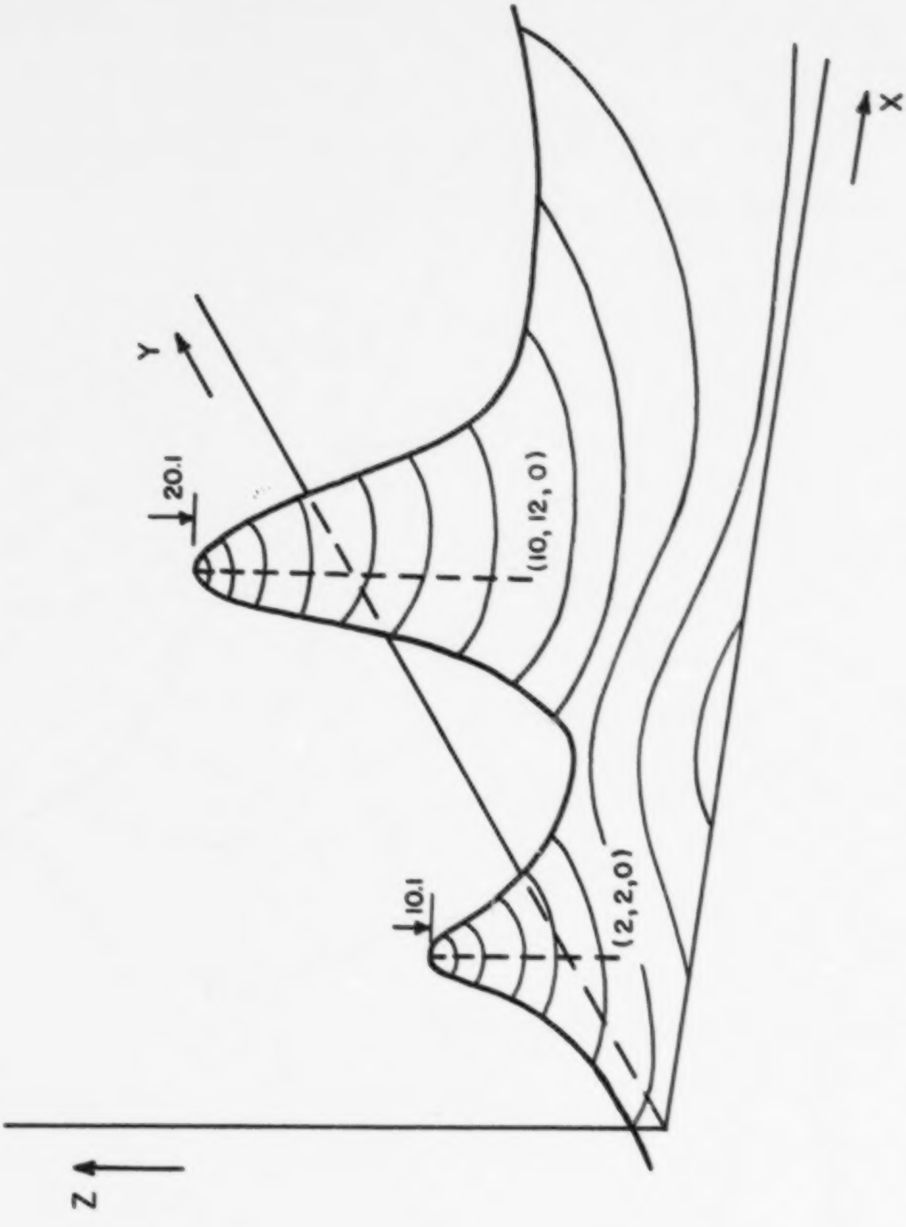
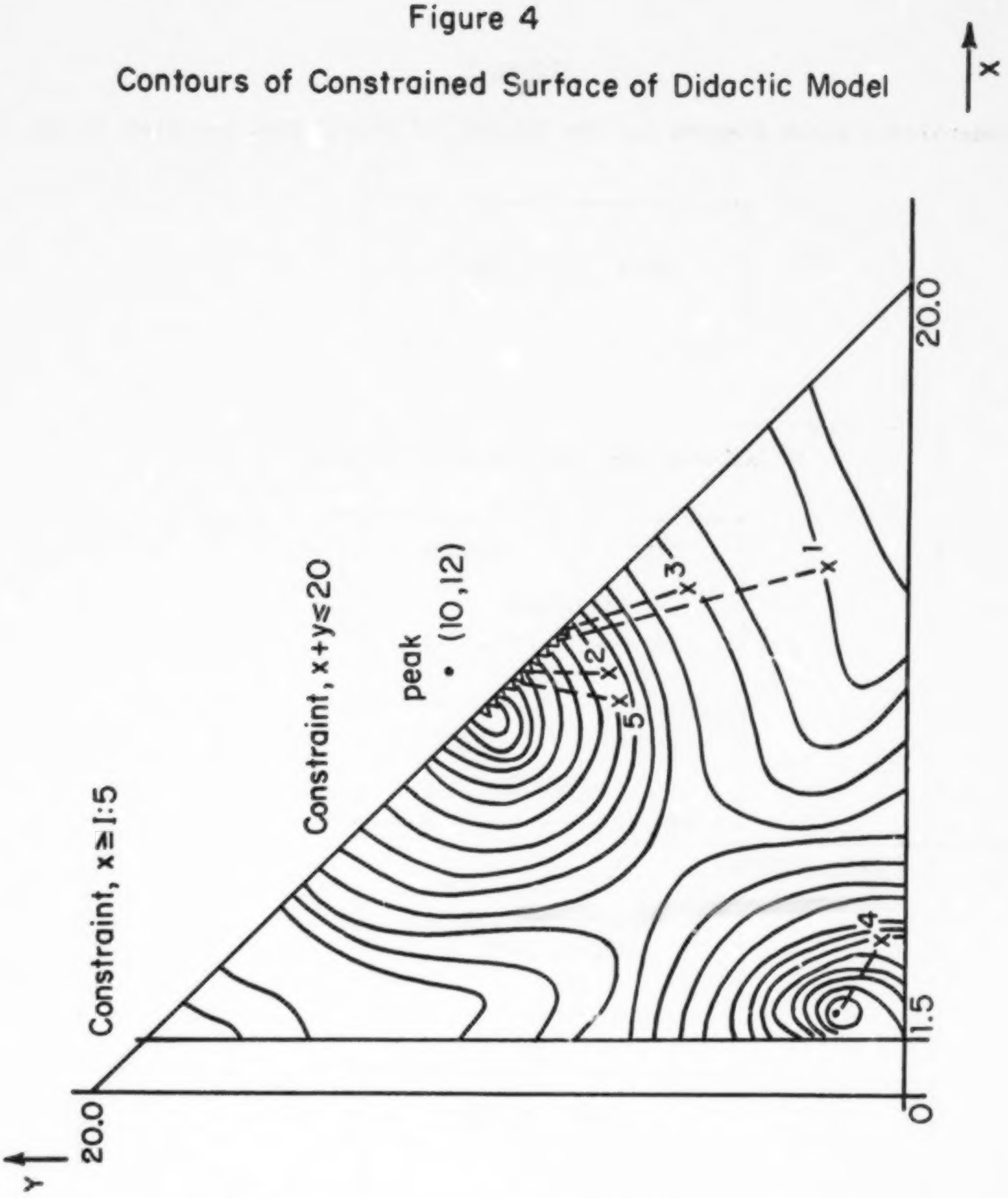

X
Y
Z
20.1
10.1
(10, 12, 0)
(2, 2, 0)

Figure 4

Contours of Constrained Surface of Didactic Model


x
Y
Constraint, x≥1.5
Constraint, x+y≤20
peak
(10,12)
20.0
1.5
0
20.0
x 1
x 2
x 3
x 4
x 5

The dotted lines represent the paths of ascent for each of the random samples taken.

FIGURE 5

Generalized Block Diagram for the Program of Constrained Steepest Ascent

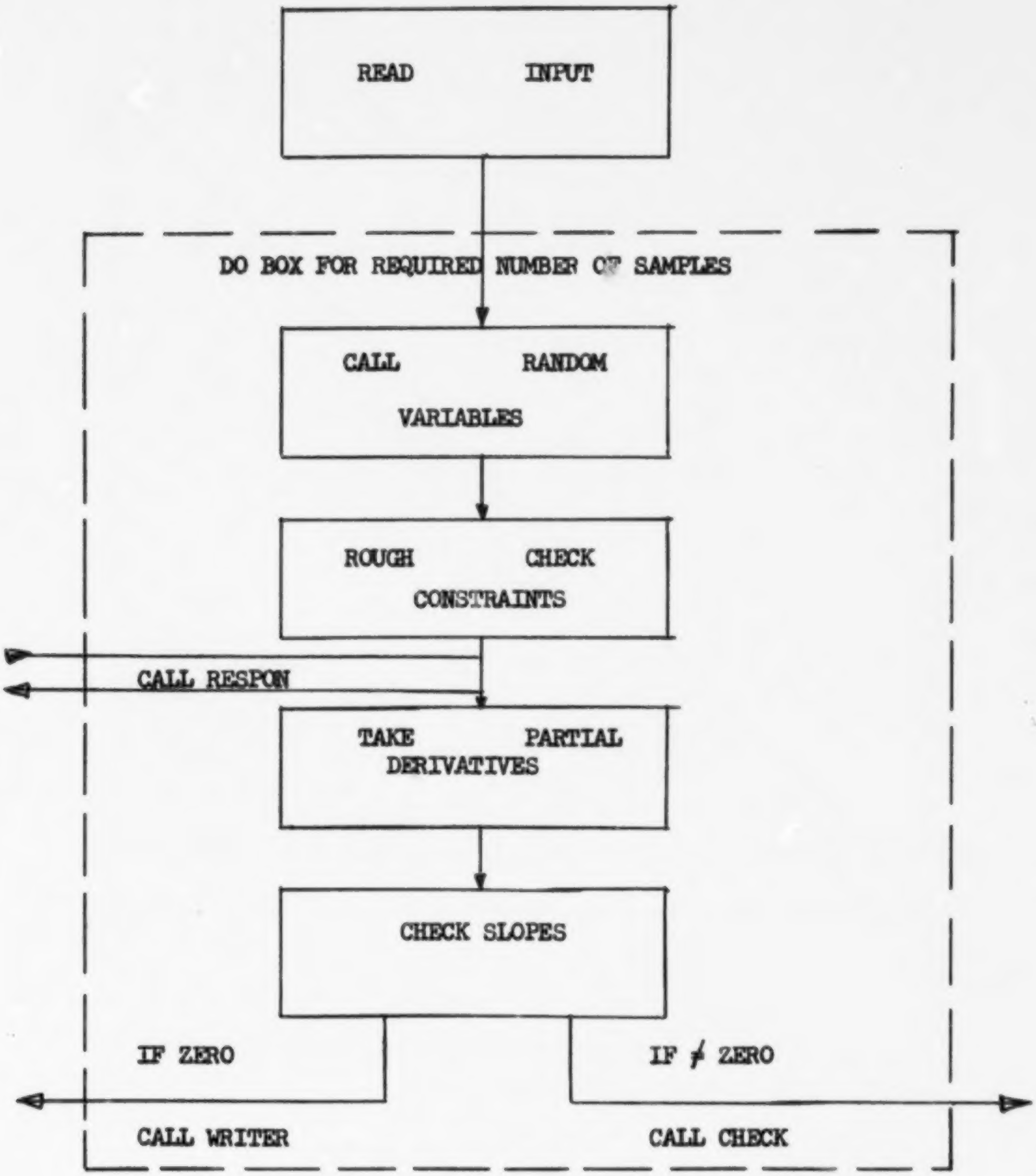
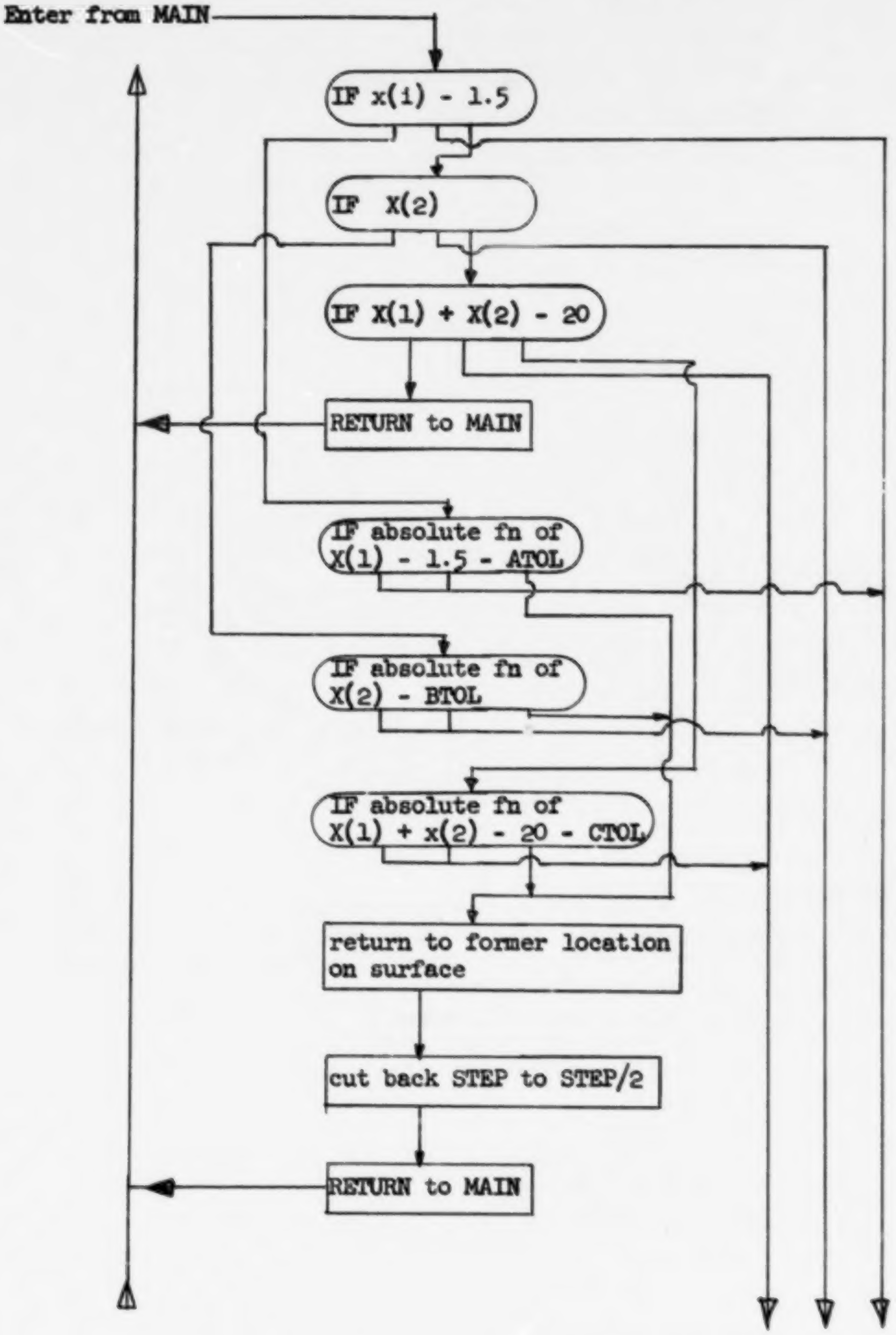

READ INPUT
DO BOX FOR REQUIRED NUMBER OF SAMPLES
CALL RANDOM VARIABLES
ROUGH CHECK CONSTRAINTS
CALL RESPON
TAKE PARTIAL DERIVATIVES
CHECK SLOPES
IF ZERO
IF ≠ ZERO
CALL WRITER
CALL CHECK

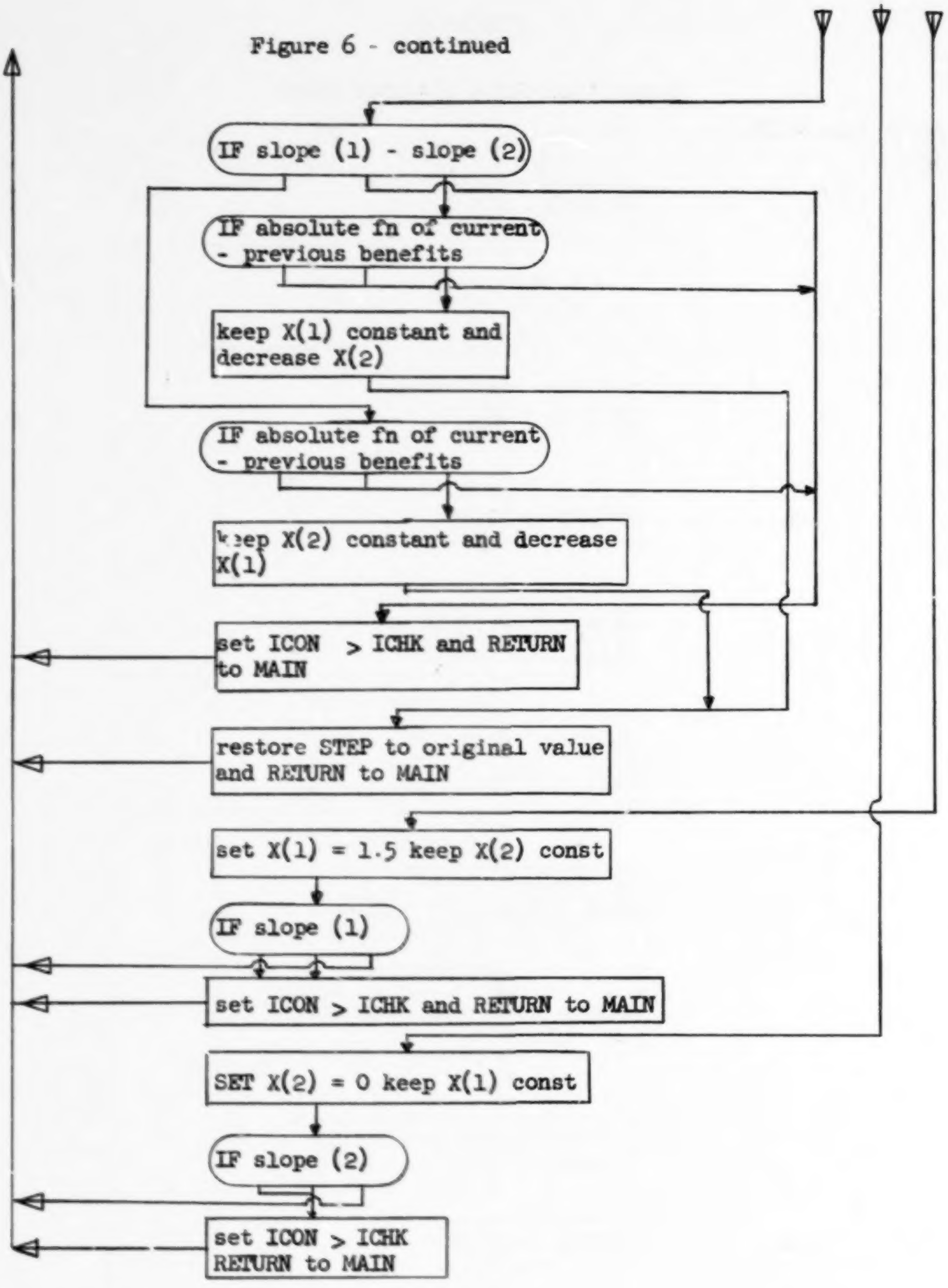
FIGURE 6

Subroutine CHECK Didactic Model

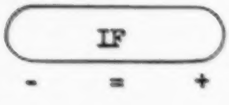

Enter from MAIN
IF x(i) - 1.5
IF X(2)
IF X(1) + X(2) - 20
RETURN to MAIN
IF absolute fn of X(1) - 1.5 - ATOL
IF absolute fn of X(2) - BTOL
IF absolute fn of X(1) + x(2) - 20 - CTOL
return to former location on surface
cut back STEP to STEP/2
RETURN to MAIN

Continued

Figure 6 - continued


IF slope (1) - slope (2)
IF absolute fn of current - previous benefits
keep X(1) constant and decrease X(2)
IF absolute fn of current - previous benefits
keep X(2) constant and decrease X(1)
set ICON > ICHK and RETURN to MAIN
restore STEP to original value and RETURN to MAIN
set X(1) = 1.5 keep X(2) const
IF slope (1)
set ICON > ICHK and RETURN to MAIN
SET X(2) = 0 keep X(1) const
IF slope (2)
set ICON > ICHK RETURN to MAIN

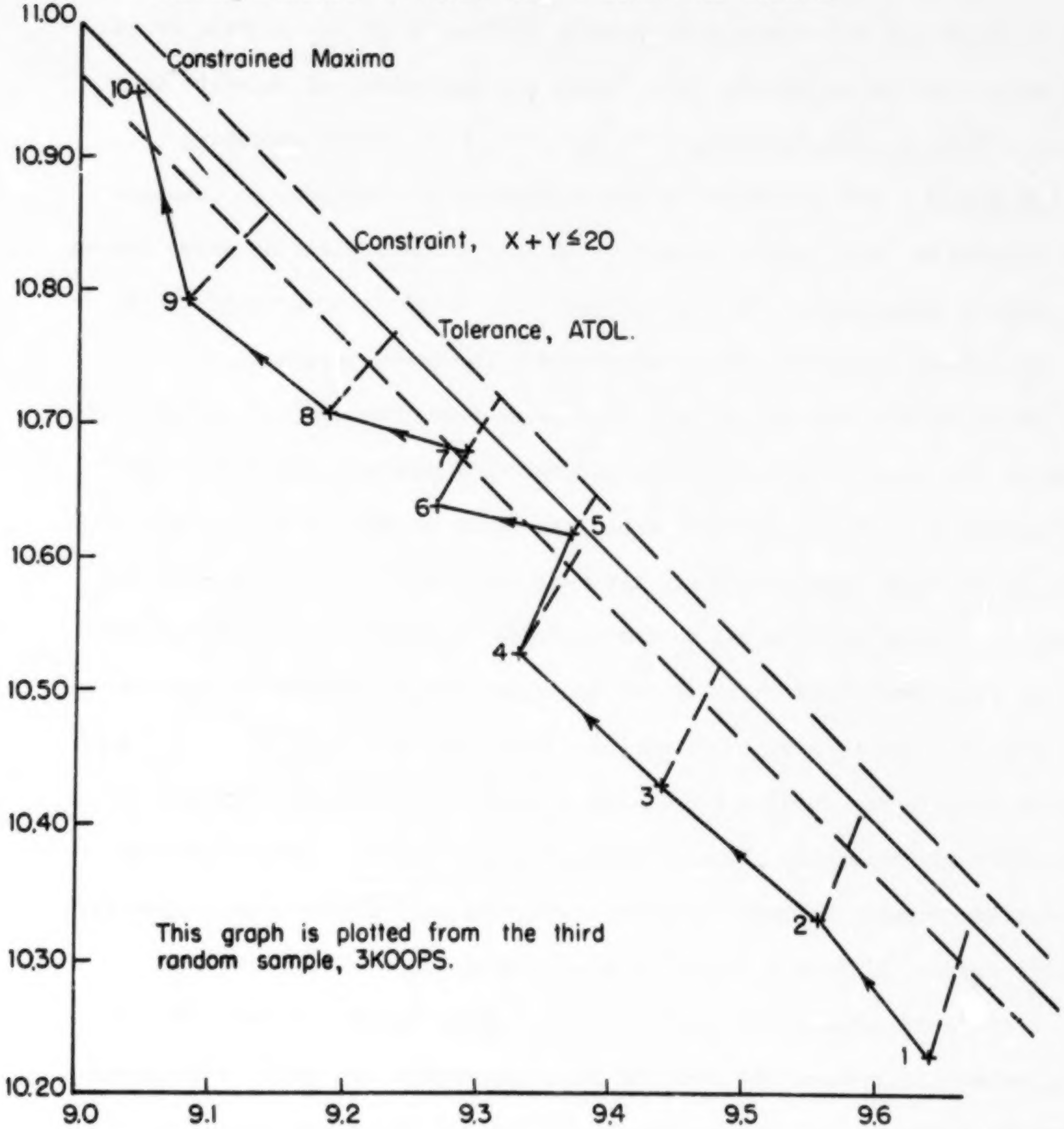
Notation:



6-21

Figure 7

Showing Behaviour of Ascent Near a Constrained Maxima



This graph is plotted from the third random sample, 3KOOPS.

6-22

CHAPTER SEVEN

Imputed Benefits of Municipal Water Supply

In designing water resource systems, the difficult problem of assigning benefits to municipal water supply is usually handled by ad hoc methods or it may even be completely by-passed. One device for assignment of benefits in common use - that of equating benefit to the cost of the least expensive alternative supply - not infrequently underestimates the benefits as measured by the criterion of "willingness to pay." In some project planning water demand is treated as a constraint. That is, a demand for water is predicated, which must be met without consideration of water needs for other purposes.

The model of this chapter shows that in a multipurpose project in which at least one of the other purposes has (or is assigned) monetary benefits, then the predication of a stated level of municipal water demand, is tantamount to the predication of an imputed benefit for municipal water. The imputed marginal benefit of the water is equal to the increase in benefits accruing to the other water uses that would occur if the municipal water requirement were reduced by one unit and this unit routed (optimally) to the other uses. It is of interest to compare the imputed benefit thus computed with prices charged for domestic water in comparable situations in adjacent regions. Also it may be of interest to compare the imputed benefit with the cost of the least expensive alternative supply, if such a supply exists. Comparisons of this sort may serve as a guide of judgment of planners in assessing whether or not the constraining water requirement has been set at a reasonable and justifiable level. Such assessment may be particularly useful in designs where the water requirement has been projected for many years into the future. In designs in which the imputed water supply benefit appears to be excessively large, maintaining the requirement at the level set may result in an uneconomic pre-emption of water from other uses.

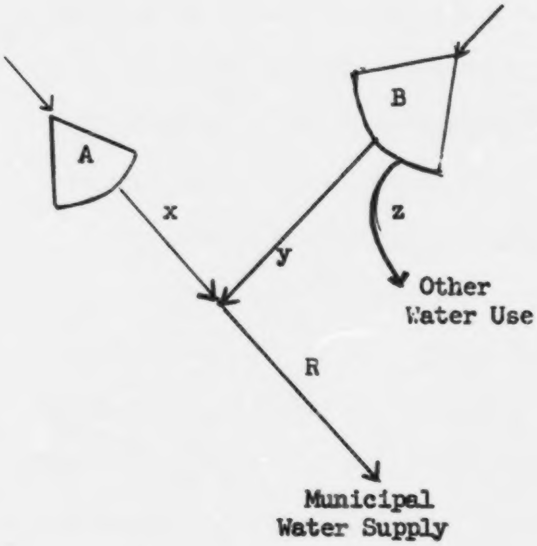
On the other hand if planners are not in a position to attach an independent rational estimate of benefits to uses such as water supply, water may in some cases be pre-empted for other uses in multipurpose projects to which explicit benefits have been assigned. This could happen if the water supply requirement had been set too low. Such pre-emption may be undesirable if benefits assigned to some of the other purposes are not directly market-connected but have been set more or less arbitrarily. If it were possible to assign a benefit figure to municipal water use rather than treating the municipal need as a constraint, the design of the most efficient system might be quite different. In this situation calculation of the imputed benefit may again provide guide to judgment.

The following simple model of a water resource system involves two reservoirs and two uses. One use is for domestic water supply for which no monetary benefit function is available but a water demand is specified; the other use is one for which there is a direct market value. The problem is (i) to determine the optimal design of the system to maximize economic efficiency subject to the constraint of the water supply requirement; and (ii) to compute the imputed benefits. The solution is obtained using the Lagrange method of undetermined multipliers.*

The two reservoirs are shown in the sketch together with outflow vectors. The outflow from Reservoir A can only be used for meeting the water supply requirement. The outflow from Reservoir B can be routed to either or both use. All flows are assumed to be non-stochastic. The following cost and benefit functions are assumed:

---

* Design of Water Resource Systems, Maass et al., Harvard University Press (1962) pp. 168-173.



The cost function for Reservoir A is $C_1 = a + x^2$, dollars, where x is the safe yield of this reservoir. The cost function for Reservoir B is $C_2 = b + (y + z)^2/2$, where y is the portion of ourflow from B routed to the municipal supply, and z the portion routed to the other use.

The benefit function for the other use is

$$B = p\,z$$

where p is the price per unit of flow. The price is assumed to be independent of flow. The municipal water requirement is

$$R = x + y$$

1). Optimum Design; both uses. The Lagrangian function to be maximized is

$$L = B(z) - C_1(x) - C_2(y + z) + \lambda\,(x + y - R) = 0 \qquad (1)$$

where λ is the Lagrangian undetermined multiplier.

By setting $\frac{\partial L}{\partial x} = \frac{\partial L}{\partial y} = \frac{\partial L}{\partial z} = 0$

it is found that

$$\frac{\partial B}{\partial z} = \frac{\partial C_1}{\partial x} = \frac{\partial C_2}{\partial y} = \frac{\partial C_2}{\partial z} = \lambda = p \qquad (2)$$

and for the optimum design

$$x = p/2, \quad y = R - p/2, \quad z = (3/2)p - R \qquad (3)$$

The net benefit will be

$$N = p\,[(3/2)p - R] - [a + (p/2)^2] - [b + p^2/2] \tag{4}$$

$$= (3/4)p^2 - pR - (a + b)$$

2). Optimum Design; water supply only. The design objective is to meet the requirement with minimum cost

$$C_T = C_1\,(x) + C_2(y)$$

$$L = C_1\,(x) + C_2\,(y) + \lambda\,(x + y - R)$$

$$\frac{\partial L}{\partial x} = \frac{\partial C_1}{\partial x} + \lambda = 0 \tag{5}$$

$$\frac{\partial L}{\partial y} = \frac{\partial C_2}{\partial y} + \lambda = 0 \tag{6}$$

From (5) and (6) $2x = y$ or $x = y/2$,

and $R = x + y = 3/2y$;

$x = R/3$; $y = 2R/3$

$$C_T = a + [R/3]^2 + b + (1/2)\,[2R/3]^2$$

$$C_T = a + b + R^2/3 \tag{7}$$

3). Optimum Design; other use only. The design objective is to maximize net benefits

$$N = B(z) - C_2(z) \tag{8}$$

$$= pz - [b + z^2/2]$$

$$\frac{dN}{dz} = p - z = 0$$

$$N = p^2 - [b + p^2/2]$$

$$= p^2/2 - b \tag{9}$$

Setting $a = 2$, $b = 3$, $p = 7$, and $R = 4$.

The optimal designs together with benefits and costs are shown in the following table

	Case	Yield A	Yield B	z	Cost	Benefit	Net Benefit
1	Both uses	3.5	7	6.5	41.75	45.5 + W	3.75 + W
2	WS only	1.33	2.67	0	10.33	W	W - 10.33
3	Other use	0	7	7	27.50	49	21.50

Imputed benefit of the municipal supply. From the table it is evident that the design of Case 1 is preferable to that of Case 2 since the net benefit is greater for all values of the water supply benefit. Hence Case 2 may be eliminated from further consideration. Now if the final decision is made to construct the design of Case 1, the lower bound for the total benefit imputed to water supply will be \$17.75 (= 21.50 - 3.75).

In a more general way we can reason as follows. If the unknown water supply benefit is included in Equation 4 and if the design of Case 1 is preferred to that of Case 3.

$$W + 0.75p^2 - pR - (a + b) \geq 0.5p^2 - b$$

$$W \geq pR - 0.25p^2 + a \qquad (10)$$

The marginal benefit of water supply

$$\frac{\partial W}{\partial R} \geq p \qquad (11)$$

Therefore for this simple water resource system the lower bound of the imputed marginal benefit is set by the price coefficient of the other use. In more complex and realistic models of proposed systems it is possible to calculate the benefits and marginal benefits imputed to water supply at any fixed requirement by an analysis similar to that used in deriving Equations (10) and (11) The imputed benefits may be compared with those from alternative supplies or from existing systems in deciding whether (or when) to build and in setting the water requirement.

7-6

GLOSSARY

autoregressive series: a series of numbers generated by a stochastic process of the following type

$$x_{t+u} = \phi(x_t, x_{t+1}, \ldots, x_{t+u-1}) + y_{t+u} ,$$

in which y is a random variable, and ϕ represents a deterministic functional relationship.

decision variable: alternatively called a "choice variable"; a system variable whose magnitude may be set arbitrarily at any desired level by the designer or planner.

deterministic process: a degenerate case of a stochastic process in which the error in forecasting the future is zero; hence in such a process a knowledge of the past is sufficient to predict the future perfectly.

economic efficiency: if the time stream of present and future costs and benefits associated with a specific project design is discounted back to the present to yield a figure representing present value of net benefit, the ratio of this figure to that corresponding to the optimal design is called the economic efficiency of the project design. In more abstract terms the economic efficiency of a water resource system is a measure of the balance of individual preferences (as manifest in "willingness to pay") for the output goods of the system against individual preferences for the goods displaced by the system.

economic time horizon: the period of time after initiation of project construction during which costs and benefits accruing to the project are discounted in evaluation of the present value of net benefit. Costs and benefits that may accrue beyond this period are neglected in the calculation of present value.

marginal benefit: the amount in dollars added to the gross benefits of a system design by the increase in one of the decision variables by one unit.

marginal cost: the amount in dollars added to the total cost of a system design by the increase of one of the decision variables by one unit (for example the increase in the cost of a water quality management system that would be entailed by increasing the capacity of one treatment plant by one million gallons per day is the marginal cost that plant).

Markov process: a stochastic process in which the conditional probability distribution for the state of the system for the next time period (or any future time) is a function only of the present state. Any additional information of past history has no relevance. Thus the outcome of any movement of the state of system depends on the outcome of the directly preceding movement, and only upon it.

mathematical model: a formalized symbolic expression or description of a casual situation which has engendered observed data.

matrix: a system of r x c numbers, real or complex, arrayed rectangularly in r rows and c columns and called an r x c matrix. An ordinary number, in contradistinction to a matrix, is called a scalar. In matrix algebra the whole matrix is treated as an entity and is subject to algebraic operations which have some resemblence to those of ordinary algebra but differ in important respects.

7-7

matrix - joint probability: a matrix the elements of which constitute the joint probability distribution of two variates.

matrix-payoff: in the theory of games, a matrix whose elements are utilities or money that passes from one person to another for all possible outcomes of a two-person game. (One of the persons may be replaced by nature as represented by a specified set of possible states).

matrix-transition: a matrix whose elements express all the conditional probabilities that obtain in the passage of the system from an initial state to a final state at some designated later time.

Monte Carlo method: a method of solution of mathematical problems, such as the solution of differential equations, that occur in a stochastic context. The meaning of the term partly overlaps that of the method of simulation.

objective function: alternatively called a "net benefit" function; it formulates the present value of the time stream of benefits and costs for alternative designs of a given project in terms of the physical decision variables (dam sizes, treatment plant capacities, etc.).

operating policy: a set of consistent and feasible rules for the management of the structural and non-structural elements of a system. A satisfactory operating procedure is one in which the given elements are manipulated to mobilize the inherent usefulness of natural water resources and to mitigate their natural destructiveness.

population: a finite or infinite collection or aggregate of individuals or items; it does not necessarily refer to a collection of living organisms.

response surface: a geometrical interpretation of an objective function in which the magnitudes of the present value of net benefit for each of all possible alternative designs of a project are conceived as forming a continuous surface in a hyperspace; the coordinates of the hyperplane are the decision variables. If there are two decision variables then the response surface for net benefit (or in special cases, minimum cost) may be represented by a contour map on a plane. All designs falling on the same contour line are equivalent in the sense that they yield the same net benefit.

season: in mathematical models a time period, such as an hour, day or month, corresponding to an aliquot of a year during which all inputs and outputs of a system flow or proceed at a constant or predetermined rate; exogeneous events can occur between but not within seasons.

serial correlation: the correlation between items of a time series and those items lagging or leading by a stated time interval. Thus for a series $x_1, x_2, \ldots$ the serial correlation of lag m is the correlation between the pairs (x_1, X_{1+m}), (x_2, x_{2+m}) etc.

simulation method: a mathematical system that is analogous to a physical system under study and which may be manipulated to produce output data closely similar to observed data. The variables of interest in the former correspond to physical variables such as flow of water or heat. Also there is a more or less complete correspondence of elements in the two systems (for example, the flow of a stream may be represented by a stochastic process).

7-8

state of nature: an expression used in game theory. The decision-maker, in playing against nature, must make a choice among a set of acts A_1, A_2 ...A_n; the relative desirability of these acts depends upon which state of nature prevails; the states of nature include all relevant and distinct courses of events that are possible between the time of decision and a specified future time at which the pay-off is made.

stochastic: as an adjective, connotes the presence of a random variable. A stochastic process is one in which at least one of the variables in the system is a random variable or has an element of randomness. Opposed to a deterministic system. The word stochastic derives from a Greek word applying to a forecast of the future course of events with the intention of predicting the truth.

stochastic variate: alternatively called "variate" or"random variable"; in contradistinction to a variable (also called "mathematical variable"), a variate is a quantity that may take any of a set of specified values (numbers or labels) in accordance with a specified set of probabilities.

target output: the physical output of a system at the breakpoint on a function relating benefits with physical output; for example, in the irrigation of a crop, increased yields or dollar return will accrue with increased application of water up to a point beyond which further applications will result in little, if any, additional yield. The amount of water per season or year at this breakpoint is the target output. The term is a misnomer in the sense that it implies it is optimal to design the system to achieve this output; in general in a multipurpose system subject to physical and budgetary constraints and in which some of the variables are stochastic variates, the optimal design is not necessarily one in which all target outputs are precisely met.

technologic function: a mathematical expression of the functional relationship between physical inputs (for example, volume of stream flow, weights of chlorine, volumes of concrete, etc.) with physical outputs (for example, volume of potable water, kilowatts of electricity, bushels of wheat, etc.).

vector: (see matrix); a matrix of r rows and one column is called a column vector; a matrix with one row and c columns is called a row vector.